AD 876563

GENERAL DYNAMICS
Electric Boat Division

PROCESSING OF DATA FROM SONAR SYSTEMS
Volume VII

This research was sponsored by the Office of Naval Research under Contract N00014-68-C-0392 (ONR Contract Authority Identification Number NR-286-001-1) with General Dynamics Corporation, Electric Boat Division. The work was performed under subcontract by Yale University.

Franz B. Tuteur
John H. Chang
Verne H. MacDonald
James P. Gray

Examined: J. W. Herring
Project Engineer

Approved: Dr. A. J. van Woerkom
Manager of Scientific Research

U417-70-051
August 11, 1970

364

ABSTRACT

Volume VII deals with the following topics:

1. Optimum Detector for Nonisotropic Noise

The previously obtained expression for the optimum detector when signal and noise are zero-mean gaussian processes, and when the noise may contain interference components are analyzed to determine the detailed structure of the detector. The detector turns out to contain beam formers that are aimed at the target signal and each interference, the signals from the interference beams being passed through rather complex filters and then subtracted from the target signal. The complexity of the optimum filter relative to conventional systems is examined, and it is found that the added complexity is quite moderate.

2. Adaptive Array Processing

The optimum detector discussed above is most easily constructed by using transversal filters, consisting of a tapped delay line and adjustable weights applied to the taps. Algorithms based on the method of stochastic approximation for automatically adjusting these weights are considered in this section and conditions for convergence and rate of convergence under several different conditions are obtained.

3. Optimum Passive Bearing Estimation in a Spatially Coherent Noise Environment

The Cramer-Rao lower bound is computed for the bearing estimator, subject to the assumption that interference noise is present. The results are compared with those obtained for a modified split-beam tracker employing simple interference nulling.

4. Space-Time Properties of Sonar Detection Models

The problem of optimizing array configurations is not a well-posed problem unless it can be shown that an optimum actually exists. Many commonly used models for sonar detection systems turn out to be singular so that the optimum does not exist i.e. it is infinite. A rigorous examination of the problem of model singularity, using measure theoretic considerations is undertaken in this section, and general criteria for nonsingularity of models are developed.

CONTENTS

FOREWORD

This is the seventh in a series of reports describing work performed by Yale University under a subcontract with Electric Boat division of General Dynamics, prime contract number N00014-68-C-0392. The Office of Naval Research is sponsor for this contract. LCDR J. F. Lyding is Project Officer for ONR. Mr. J. W. Herring is Project Engineer for Electric Boat division under the direction of Dr. A. J. van Woerkom, Manager of Scientific Research.

I. Introduction

This report is the first of two volumes dealing with work completed under contract 8050-31-55001 between Yale University and the Electric Boat Company during the period from July 1, 1968 to April 30, 1970. More detailed discussions of the results are contained in the four progress reports Nos. 38, 39, 40 and 41, which are appended. The companion volume (vol. VIII of this series) covers work done during the same time period and contains results submitted originally in progress reports No. 42 and 43. Three of the topics contained in this volume are continuations of work covered in earlier reports, dealing with the effects of anisotropy of the background noise field - also referred to as interference noise. The three progress reports deal respectively with the form of the optimum detector, with the behavior of adaptive detectors, and with bearing estimation under these noise conditions. The fourth topic, which deals with the effect of signal models and the various possibilities for singular detection is entirely new and represents a substantial departure from work described in previous reports.

II. The Optimum Detector for Nonisotropic Noise

An expression for the optimum detector transfer function when the noise contains one or more strong interference components was originally obtained in Progress report No. 33, which is part of volume V of this series. The implication of this expression on the detailed structure of the detector is examined in Progress report No. 38.

The results of both reports are based on the assumptions that the signal, noise, and interference are all sample functions of a zero-mean gaussian random process, that the interference consists of a number of isolated point sources and that the noise is otherwise isotropic and far-

field. Under these conditions the filter can be shown to separate into a spatial part - essentially a set of beam formers - and a temporal part or Eckart filter. For the case of a single interference the spatial part, which is also the significant part, takes the form:

$$H''(\omega) = \begin{bmatrix} e^{-j\omega\tau_1} \\ \vdots \\ e^{j\omega\tau_M} \end{bmatrix} - \frac{K_1(\omega)G_{10}(\omega)}{1 + MK_1(\omega)} \begin{bmatrix} e^{-j\omega\tau_1^{(1)}} \\ \vdots \\ e^{-j\omega\tau_M^{(1)}} \end{bmatrix}$$

where the τ_i are the signal delays, the $\tau_i^{(1)}$ are the interference delays, M is the number of hydrophones, $K_1(\omega)$ is the ratio of interference spectral density to ambient noise density, and $G_{10}(\omega)$ is given by

$$G_{10}(\omega) = \sum_{k=1}^{M} e^{-j\omega(\tau_k - \tau_k^{(1)})}$$

This result can be interpreted to mean that the filter contains a simple beamformer aimed at the signal and a second beamformer aimed at the interference, and that the interference output is subtracted from the signal output after being passed through a filter with the transfer function given by the coefficient of the second bracketed term in the above expression.

For more than one interference the result is basically similar - a beam is aimed at each interference and the output is subtracted from that of the main signal beam after passage through a compensating filter. The complexity of these compensating filters increases with the number of the interferences; in fact even for a single interference it is such that automatic design by some sort of adaptive mechanism would almost have to be used. This point is considered further below.

A major difficulty in the design is that because of the need to form several beams simultaneously, beam steering must be done by tapped delay

lines. The number of taps and the tap spacing are largely a function of array resolution, which in turn, can be related to array aperture. For typical arrays the number of taps tends to be very large; however this is true even if conventional or suboptimal instrumentations are used. The added complexity required in the optimal instrumentation is, from this point of view, quite modest.

III. Adaptive Array Processing

The automatic design of complicated filter transfer functions of the sort mentioned above can be accomplished fairly easily by means of transversal filters - filters produced by feeding a signal into a tapped delay line and adding the weighted tap outputs to form the output. For a delay line having M taps the output y(t) of such a filter has the form

$$y(t) = \sum_{i=1}^{M} c_i \, x(t-\tau_i)$$

where x(t) is the input signal, and c_i is the weight applied to the intput delayed by the time τ_i. If $\tau_i - \tau_{i-1}$ is small for all i = 1...M this expression is a discrete approximation of a convolution integral in which the c_i represent the impulse response of a filter at time t_i. Since each c_i can take on any arbitrary value, extremely complex filters are easily synthesized in this way. It was shown in progress report No. 34 that the adjustment of the c_i subject to one of several criteria of optimality is easily accomplished by means of algorithms based on the stochastic approximation method of Robbins and Monro. Progress report No. 39 is a continuation and elaboration of the earlier report.

The basic assumptions used in the analysis are:

1) Target, interference, and ambient noise are zero mean gaussian processes
2) The sum of interferences, ambient noise, and local noise are regarded as

the effective noise, which is assumed to be statistically independent of the target signal.

3) The target signal component $s_i(t)$ observed at the output of the i^{th} hydrophone is a linear time invariant transformation of d(t), the target-signal that would be observed at the output of an ideal isotropic hydrophone located at the origin of coordinates. The autocorrelation function of d(t) is assumed to be known.

4) The statistics of the noise field are unknown. It is not known whether interferences are present, or where they are located.

5) The wave fronts of target and interference are assumed to be plane over the dimensions of the receiving array.

It is assumed that the adaptive mechanism is to produce a filter optimized in a given direction and designed to suppress interference signals from other directions. By varying the azimuth for which the filter is optimized the system produces a bearing response pattern which can be examined by an operator to determine whether a target is present.

The space-time filter takes the form of a set of K hydrophones, each connected to the input of the delay line of a transversal filter having M taps. (Note that this notation differs from that used in most of the other reports in this series). The outputs of all the transversal filters is summed to form the signal z(t), which after possible further filtering, is squared and smoothed to yield the observed output. The adjusting algorithm for the K(M+1) weights in all of the transversal filters then takes the simple form:

$$\underline{W}_{j+1} = \underline{W}_j + 2\gamma_j \, [\underline{R}_{d\xi} - z_j \, \underline{\eta}_j]$$

where $\underline{W}_j$ is the vector of all of the tap weights suitably indexed, γ_j is a weighting parameter, $\underline{R}_{d\xi}$ is the input space-time autocorrelation function,

$\underline{z}_j$ is the output z(t), and $\underline{\eta}_j$ is a vector of delayed versions of the received signal; all at the j^{th} step in the iteration. The process converges for γ_j of the form $\gamma_j = \gamma/j^{\alpha}$ with $\frac{1}{2} < \alpha \leq 1$. $\underline{R}_{d\xi}$ can be computed if the target signal direction and autocorrelation function are known; thus it contains the information about desired target direction that is needed for the filter to adjust itself.

General expressions for the convergence of the filter have been obtained and are given by Eqs. 3.5 - 32 and 3.5 - 33 of progress report 39. These expressions are too complicated to yield much insight. They can however be simplified by chosing specific expressions for the weighting parameter γ_j. A particularly simple expression results from the choice $\gamma_j^{(m)} = \frac{1}{2(j+1)\lambda_m}$, where λ_m is the m^{th} eigenvalue of the covariance matrix of the received signal, and where the superscript (m) on γ_j implies that different weights are used in different filters. In this case it is found that the mean-square error at the $(j+1)^{th}$ step is given by

$$\overline{e_{j+1}^2} - \overline{e_{min}^2} \approx \frac{j}{(j+1)^2} K(M+1)\, \overline{e_{min}^2} + \frac{1}{(j+1)^2} (\underline{W}_1 - \underline{W}_{op})^T R_y (\underline{W}_1 - \underline{W}_{op})$$

where R_y is the covariance matrix of the received signal, and where e_{min} is the irreducible error resulting from the fact that a continuous filter is approximated by a discrete structure. If the second term is initially larger than the first then this expression indicates an initial m.s. error reduction at a rate j^{-2}; however eventually the first term will always dominate, with the result that convergence eventually takes place at a rate j^{-1}.

As long as the noise environment is stationary the filter converges to the optimum form discussed in previous progress reports (e.g. #38) in which the interference noises are strongly suppressed. This is shown not

only analytically, but also by means of a computer simulation using real data. If the noise environment is nonstationary, partial results have been obtained under the following conditions:

1. If the nonstationarity can be characterized by changing parameters, with the values of the parameters governed by a known dynamic relation than the method of stochastic approximation can be modified by inclusion of this dynamic relation. In fact the recursive Kalman filter method can be applied to this case with results that converge to those obtained by the method of stochastic approximation in the stationary case. In the nonstationary case the weighting parameter γ_j of the stochastic approximation algorithm is modified and takes the form $\gamma_j' = \gamma_j + \beta$, where β is a constant. For the case where the optimum gain parameter $\underline{\theta}_j$ is given by the relation

$$\underline{\theta}_{j+1} = a\,\underline{\theta}_j + u_j, \qquad 0 < a < 1$$

and where the desired filter output is given by

$$D_j = \underline{H}_j\,\underline{\theta}_j + v_j$$

where

$$H_j = \begin{bmatrix} \underline{n}_1^T \\ \vdots \\ \underline{n}_j^T \end{bmatrix}$$

and where u_j and v_j are stationary independent, zero mean, scalar, white noise processes, with variances q and ϕ respectively, then

$$\beta = q/\phi$$

For the stationary case $q = 0$ and $a = 1$, so that $\beta = 0$, but in general the presence of a nonzero β prevents the gradual disappearance of the weighting parameter γ_j, which would make tracking of a changing environment impossible. On the other hand, the fact that γ_j does not go to zero

as $j \to \infty$ has the effect that the filter does not converge in mean square, which means that a small jitter (proportional to β) continues to exist in the output.

2. If the nonstationarity is such that the optimum gain parameter $\underline{\theta}_j$ satisfies a relation of the form

$$\underline{\theta}_{j+1} = \underline{\theta}_j + 0\left(\frac{1}{j^s}\right)$$

i.e. the nonstationary is in a sense "temporary" and disappears with $j \to \infty$, then the standard method of stochastic approximation converges as long as the weighting factor γ_j has the form

$$\gamma_j = \frac{\gamma_0}{j^\alpha}$$

where $\frac{1}{2} < \alpha \leq 1$

and $s < \alpha$

Other methods for dealing with nonstationary environments can be envisioned, but have not yet been evaluated.

IV. Optimum Passive Bearing Estimation in a Spatially Coherent Noise Environment

Report No. 40 is a continuation of report No. 37 which was included in vol. V. The earlier report dealt with the Cramer-Rao lower bound for determining the rms bearing error attainable in an isotropic noise field. The present report extends this to the case where interference is present. As in the earlier report the analysis initially considers an arbitrary number of hydrophones arbitrarily spaced on a linear array, and arbitrary signal, ambient noise, and interference spectra. However, in order to get results that are simple enough to yield some insight into important parameters, some of this generality is sacrified; in particular it is

assumed that the ambient noise power is much greater than the signal power; signal, interference, and ambient noise spectra are taken to be identical in form, and the hydrophone spacing is uniform. Additional important assumptions are that the interference bearing is known and that the ambient noise is independent from hydrophone to hydrophone. Also, as in the earlier report the performance of the split-beam tracker is computed to provide a comparison between the possibly unrealizable bound and a practical instrumentation.

Approximate expressions for the lower bound take on simple forms if the target and interference separation is either very large or very small. In each case limiting expressions have been obtained for the ambient-noise dominated case ($MI \ll N$) and for the interference dominated case ($MI \gg N$). The parameter determining target and interference separation is $y = (d/c) \cdot \omega_{max} (\sin\theta - \sin\theta)$ where d is the hydrophone spacing, c is the sound velocity ω_{max} is the maximum frequency, θ is the target bearing and ϕ is the interference bearing. If the bias terms are neglected then for $y \gg 1$ the respective lower bounds are approximately

$$\overline{(\hat{\theta} - \bar{\hat{\theta}})^2} \geq \begin{cases} \dfrac{36\pi c^2 (N^2 + MNS)/S^2}{T\,\omega_{max}^3\, d^2 \cos^2\theta (M^4 - M^2)[1 + (M-2)I/N]} & (MI \ll N) \\[2ex] \dfrac{36\pi c^2 [N^2 + (M-1)NS]/S^2}{T\,\omega_{max}^3\, d^2 \cos^2\theta [M^4 - M^2 - 8/5\, M^3 + 2M]} & (MI \gg N) \end{cases}$$

The lower bound for $I = 0$ (i.e. no interference) is the same as that found in report No. 37. By comparing the denominators in the two expressions above one can conclude that the effect of a remote interference is equivalent to the loss of 2/5 of a hydrophone.

For near interference, such that $y < 1/M$, the corresponding results are

$$\overline{(\hat{\theta} - \bar{\theta})^2} = \begin{cases} \dfrac{36\pi c^2 (N^2+MSN)/S^2}{T\,\omega_{max}^3\, d^2\cos^2\theta (M^4-M^2)[1+(M^4-M^2)I^2y^2/(10N^2)]} & (MI \ll N) \\[2ex] \dfrac{36\pi c^2 MI[N+(M-1)S]/S^2}{T\,\omega_{max}^2\, d^2\cos^2\theta (M^4-M^2)} & \begin{pmatrix} MI \gg N \\ y \ll \dfrac{3N}{M^2 I} \end{pmatrix} \end{cases}$$

If the difference in denominators (which amounts to the previously mentioned 2/5 M) is discounted, the lower bound in the interference dominated case is seen to be MI/N times as large as for large y.

A modified form of split-beam tracker employing simple interference nulling ahead of the split-beam section was considered in progress report No. 29. For this tracker the following results were obtained:

$$\overline{(\hat{\theta} - \bar{\theta})^2} = \begin{cases} \dfrac{96\pi c^2 N^2/S^2}{T\,\omega_{max}^3\, d^2\cos^2\theta\, M^2(M-2)^2} & \begin{pmatrix} y \gg 1 \\ M \gg 1 \end{pmatrix} \\[2ex] \dfrac{(64)^2 \pi c^2 N^2/S^2}{T\,\omega_{max}^3\, d^2\cos^2\theta\, M^2(M-2)^2} \cdot \dfrac{1}{y^4} & \begin{pmatrix} y \ll 1 \\ M \ll 1 \end{pmatrix} \end{cases}$$

The second of these expressions is invalid for y very near zero because some of the approximations made to obtain it break down, however it is an indication that the modified split-beam tracker cannot estimate bearings extremely close to the interference bearing (since as a result of the nulling there is no signal in this direction). In this respect the split-beam tracker performance appears to fall considerable short of the Cramer-Rao bound, which is finite for y = 0, albeit consideralby larger than for large separation. The comparison between the split beam tracker and the Cramer-Rao bound is facilitated by computing the ratio of the two error variances, this is

$$\frac{\sigma^2_{sbt}}{\sigma_{cr}} = \begin{cases} 2.67(1+2.4/M) & y >> 1, \quad M >> 1 \\ \frac{11.4}{y^2}(1+8/M) & 0 < y << 1, \quad M >> 1 \end{cases}$$

The second of these expressions is invalid for $y = 0$ as indicated above. Both expressions indicate that for sufficiently large M the split-beam tracker performance is fairly close to the lower bound; but at the same time they also suggest that some improvement might be achieved, particularly for small separations between target and interference, by going to a different implementation. Such implementations are currently being studied.

By plotting curves for the exact expressions relating $\overline{(\hat{\theta} - \bar{\theta})^2}$ to y it is found that the large y approximation is good for separations between target and interference bearing greater than the beam width of the array, defined as the angle for which the signal output falls to one half its maximum value. This is roughly true both for the C-R bound and the split-beam tracker. Also, in both cases, for separation smaller than the beam-width the performance deteriorates rapidly; however the deterioration is considerably more rapid in the case of the split-beam tracker.

The error variance decreases with M^4 for large separations in both cases, and the C-R bound decreases with M^3 for zero separation between target and interference bearing. Thus theoretically the error can be made arbitrarily small for both large and small separations by letting M become sufficiently large. Here it must be noted however, that for a fixed size array the assumption of zero ambient noise correlation between adjacent hydrophones will become invalid for very large M.

V. Space Time Properties of Sonar Detection Models

In all previous work, and in most analyses of sonar in the literature the array configurations are taken as given. In a good many of the analyses

reported in the previous volumes, in fact, the arrays have been assumed to be linear and with equally spaced hydrophones. The question naturally arises as to whether the performance of an array with a given number of hydrophones might not be improved substantially by seeking an optimum configuration.

It turns out that the attempt to find algorithms for determining the optimum placement of hydrophones involves searches through a 3K-dimensional continuum, where K is the number of hydrophones. For the large values of K that are of practical interest such a search is an extremely formidable undertaking for which there is no guarantee of success. Hence it becomes very desirable to obtain first some estimate for the ultimate performance of which an array with a large number of arbitrarily spaced hydrophones is capable. Such an estimate is, however, even conceptually possible only if in the limit of continuous observation (i.e. as $K \to \infty$) the signal model remains nonsingular. Many commonly used models turn out, in fact, to be singular; i.e. as $K \to \infty$ it becomes possible to determine the presence or absence of the signal with zero error even though both the array size and observation times are finite. For this reason such models are physically not completely realistic (which is not to say that they are not useful), and it is desirable to obtain general conditions guaranteeing that a given model be nonsingular. This is in essence what is done in report #41.

The approach taken is based on the realization that any communication/detection (C/D) system can be represented as a series of mapping operations; i.e. an encode operator e maps source characters from the space A of source characters into the space W of channel signals which is in turn mapped by a transmit operator t into a space V of receivable signals, etc. until the final mapping produces an estimate $\hat{a}$ of the source character, which is an element of the space $\hat{A}$. The operators are stochastically determined, hence

can be considered as being themselves elements of probability spaces E, T, etc. The notation is generalized by denoting the space of source characters by S_1 and the space of mappings of S_{2K-1} into S_{2K+1} by S_{2K} where $K = 1,2,3...L$. A marginal probability measure μ_i may then be defined on each of the spaces S_i where $i = 1,2,4,6...2L$. These measures will induce probability measures in the remaining spaces S_{2K+1}, $K = 1,2,3...L$; furthermore they induce conditional measures of the form μ_i^a, the measure induced in S_i conditioned on the transmission of a signal a.

A class of models having particularly simple properties are the factorable models. In this class the probability measure μ defined on the product space $S = S_1 \times S_2 \times S_4 \ldots S_{2L}$ is given by

$$\mu = \mu_1 \mu_2 \mu_4 \cdots \mu_{2L}$$

This form of the measure implies that the stochastic operations of the model are independent. Most of the models used in the usual communication and detection studies are factorable in this sense.

The central theoremes concerning the singularity of models are then given by Corollary 1 of Theorem 9, and Theorem 10 of chapter 2:

If $S_1 = S_{2L+1}$ are countable with discrete metric, then the model M is singular if and only if the conditional measures μ^r_{2L+1} and μ^s_{2L+1} are orthogonal for every pair of characters r and s in S_1 for which $\mu(r) > 0$ and such that $r \neq s$.

If M is a factorable model then it is singular if and only if μ^r_{2k+1} and μ^s_{2k+1} are orthogonal for all $k < L$, and $r \neq s$.

The implication of these theorems is that for a factorable model to be singular, singularity most be present in the first stage, and it must be preserved by all subsequent transformations or mappings. While this might appear to be a rather strong requirement which would have the effect

of making most practical models nonsingular, it turns out that many of the usual encoding transformations considered in communications processes preserve singularity, so that singular models are actually more common than might be supposed. In particular, it is shown in Theorem 1 of chapter 3 that additive stages are usually singularity preserving. On the other hand it is shown in Theorem 2 of this chapter that if stage k is such that the support space of the measure μ_{2k} is a subspace of the previous space (S_{2k-1}) and if μ_{2k} is independent of μ_i for all $i < 2k$ then the model M is nonsingular.

Particularly simple statements can be made if the conditional measures μ^r_{2k+1} are Gaussian. In this case one can use the fact that two Gaussian measures are either orthogonal, or they are equivalent. Furthermore according to Theorem 4 of chapter 3 two Gaussian distributions P and Q are equivalent if and only if

1. $m(\cdot) \in H(\Gamma_Q)$
2. Γ_P has a representation $\Gamma_P(s,t) = \Sigma\, \lambda_k\, e_k(s)\, e_k(t)$ where the set of functions $e_k(t)$ is a complete orthonormal set in the reproducing kernel Hilbert space $H(\Gamma_Q)$ and $\Sigma(1-\lambda_k)^2 < \infty$, and $\lambda_k > c > 0$ for all k.

In this theorem Γ_P and Γ_Q are covariance of the distributions P and Q, with mean functions $m(\cdot)$ and 0 respectively. As a consequence of this theorem singularity may occur when the mean function of the signal process lies outside the space $H(\Gamma_Q)$; i.e. if P has a linear projection outside the support space of Q; if this is not the case singularity may still occur if some noise eigenvalues are zero or if the signal and noise processes do not put almost the same energy into all but a finite number of dimensions (or eigenvalues).

Applications of this theory have been made to two simple sonar situations. The first of these is one dimensional: a source is either to the right or to the left of the observer, and the observer can determine the direction of wave propagation. This situation is singular, even if the velocity of propagation

is random, if random noise is added, and if other random effects are present, as long the randomness is not sufficient to make a right-going wave look like a left going wave.

The more interesting problem of sonar in three-dimensions has also been analyzed with the result that the usual model in which the signal wavefront isa deterministic function of the coordinates is also shown to be singular. This explains the result of Vanderkulk that as the number of hydrophones goes to infinity, the array gain becomes infinite and detection becomes perfect. This result is shown to hold even if white noise is added at each hydrophone; to produce a nonsingular model it is necessary to introduce some perturbations into the wavefront. The effect of perturbed wavefronts is currently being analyzed.

THE OPTIMUM DETECTOR FOR NONISOTROPIC NOISE

by

Franz B. Tuteur

Progress Report No. 38

General Dynamics/Electric Boat Research

September 1968

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY

Summary

The feasibility of using tapped-delay-line filters to synthesize the optimum processor is investigated in this report. It is found that the most severe requirement that is placed on the delay lines arises from the necessity of steering the array in steps that are commensurate with the resolution of which the array is theoretically capable. If delay lines to accomplish this can be fabricated then the additional complexity required in the construction of an optimal filter is relatively minor; that is, it requires delay lines of no greater complexity.

The Optimum Detector for Nonisotropic Noise

I. Introduction.

In Progress Report No. 33 (Ref.1) the effect of localized noise sources on the performance of the optimum (likelihood-ratio) detector of directional Gaussian signals was investigated. In the present report the structure of the optimum detector is considered.

The nomenclature used in this progress report is exactly the same as that used in Ref. 1, which is assumed to be available to the reader.

II. General Form of the Optimum Detector

The general form of the optimum detector is contained in Eqs.(22) and (23) of Ref. 1. If the output of the filter is designated by

$$u = \log LR - C \tag{1}$$

then the optimum detector structure has the form

$$u = \sum_{n=1}^{WT} |\underline{H}^T(n)\underline{X}(n)|^2 \tag{2}$$

where $$\underline{H}(n) = \frac{\sqrt{S(n)}}{N(n)} \frac{\underline{Q}^{-1}(n)\ \underline{V}^*(n)}{\sqrt{1 + S(n)G_o(n)/N(n)}} \tag{3}$$

and where the optimum array gain $G_o(n)$ is defined by

$$G_o(n) = \underline{V}^T(n)\underline{Q}^{-1}(n)V^*(n) \tag{4}$$

If the time of observation T is large, the summation in (2) can be converted to an integral in the frequency variable f; hence the detector output u takes the form

$$u \approx T \int_0^{\omega} |\underline{H}^T(f)\underline{X}(f)|^2 df \tag{5}$$

where by direct analogy with Eq.(3)

$$\underline{H}(f) = \frac{\sqrt{S(f)}}{N(f)} \frac{\underline{Q}^{-1}(f)\ \underline{V}^*(f)}{\sqrt{1 + S(f)G_o(f)/N(f)}} \tag{6}$$

In this expression $S(f)$ and $N(f)$ are, respectively the signal and noise spectral densities, and $\underline{Q}(f)$ is the noise spectral matrix whose elements are the cross spectral densities of the noise voltages received on the different hydrophones of the array. $\underline{V}(f)$ is the steering vector, given by

$$\underline{V}(f) = \begin{vmatrix} c_1 e^{j2\pi f\tau_1} \\ \vdots \\ c_M e^{j2\pi f\tau_M} \end{vmatrix} \tag{7}$$

where the c_i are hydrophone gains and τ_i the signal delays. Also, the array gain becomes

$$G_o(f) = \underline{V}^T(f)\underline{Q}^{-1}(f)\underline{V}^*(f) \tag{8}$$

If the bandwidth W is very large, Parseval's theorem can be used to convert Eq.(5) into

$$u \simeq W \int_0^T |z(t)|^2 dt \tag{9}$$

where $z(t)$ is the inverse Fourier transform of $\underline{H}^T(f)\underline{X}(f)$. This implies that u can be obtained from a circuit of the form shown in Fig. 1.

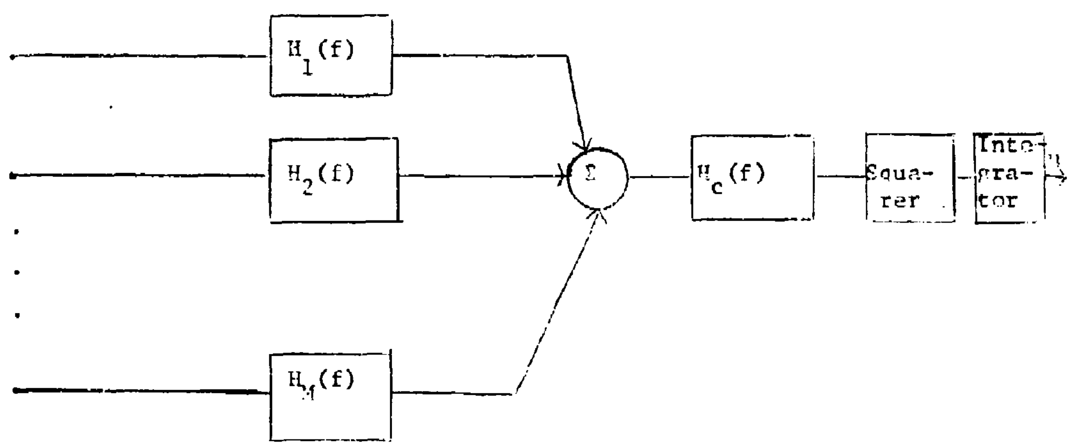


Figure 1. Likelihood-ratio Detector

In this figure $H_c(f)$ is a filter containing the common frequency-sensitive component in $\underline{H}(f)$, namely

$$H_c(f) = \frac{\sqrt{S(f)}}{N(f)\sqrt{1+S(f)C_o(f)/N(f)}} \tag{10}$$

The individual filters $H_1(f)$, $H_2(f)\ldots H_m(f)$ are then respectively the first, second, ... , Mth row of the matrix product $\underline{Q}^{-1}(f)\underline{V}^*(f)$.

III. Detailed Structure of the Filter with Directional Interference.

As in section IV of Ref. 1 we assume that the noise component consists of an isotropic part and a number of point sources. Then the noise spectral matrix has the form

$$\underline{Q}(f) = \frac{N_o(f)}{N(f)} [\underline{Q}_o(f) + \sum_{r=1}^{R} K_r(f)\underline{V}_r^*(f)\underline{V}_r^T(f)] \tag{11}$$

where $K_r(f) = \frac{I_r(f)}{N_o(f)}$ and where $I_r(f)$ is the spectral density of the r^{th} noise source.

Since the frequency weighting filter of Eq.(10) is common to all channels, the essential operation performed by the processor is

$$\underline{H}'(f) = \underline{Q}^{-1}(f)\underline{V}^*(f) \tag{12}$$

The inversion of the spectral matrix in (12) can be accomplished by means of Eq.(35) in Ref. 1. The result is

$$\underline{H}'(f) = \frac{N(f)}{N_o(f)} \{\underline{Q}_o^{-1}\underline{V}_o^* - \underline{Q}_o^{-1}[\sqrt{K_1}\underline{V}_1^* \vdots \sqrt{K_2}\underline{V}_2^* \vdots \ldots \vdots \sqrt{K_R}\underline{V}_R^*][\underline{I}+\underline{G}]^{-1}\underline{g}^*\} \tag{13}$$

where the dependence of $\underline{Q}_o$, $\underline{V}_r$, K_r, $\underline{G}$, and $\underline{g}$ on f has been suppressed for convenience. The notation of Ref. 1 is used, with n replaced by f in all cases. Note that the scalar multiplying factor $N(f)/N_o(f)$ can be absorbed into $H_c(f)$ of Eq.(10) so that the essential operation is that indicated by the expression in the braces, {...}.

To gain some insight into the implications of this result we consider some simple examples. In all cases we assume that $\underline{Q}_o(f) = I$, the unit matrix; this implies that there is no interphone correlation of the isotropic noise component.

Suppose first that there is only a single interference. Then the expression inside the braces in Eq.(13) becomes

$$\underline{H}''(f) = \frac{N_o(f)}{N(f)}\,\underline{H}'(f) = \underline{V}_o^* - \frac{K_1G_{10}}{1+K_1M}\,\underline{V}_1^*$$

$$= \begin{bmatrix} e^{-j\omega\tau_1} \\ \vdots \\ e^{-j\omega\tau_M} \end{bmatrix} - \frac{K_1G_{10}}{1+K_1M} \begin{bmatrix} e^{-j\omega\tau_1^{(1)}} \\ \vdots \\ e^{-j\omega\tau_M^{(1)}} \end{bmatrix} \tag{14}$$

where $\omega = 2\pi f$ and where

$$G_{10} = G_{10}(f) = \sum_{k=1}^{M} e^{j2\pi f(\tau_k^{(1)} - \tau_k)} \tag{15}$$

The unsuperscripted τ's are the signal delays, while the superscripted τ's are the interference delays. Thus Eq.(14) indicates that the filter forms two beams, one steered on the signal and one on the interference, the output of the interference beam is passed through a filter with transfer function $K_1G_{10}/(1 + K_1M)$ and the result is subtracted from the signal beam output. A possible system block diagram is shown in Figure 2. This system is quite similar to that proposed by V.C. Anderson [4] and reported on by Cox [5].

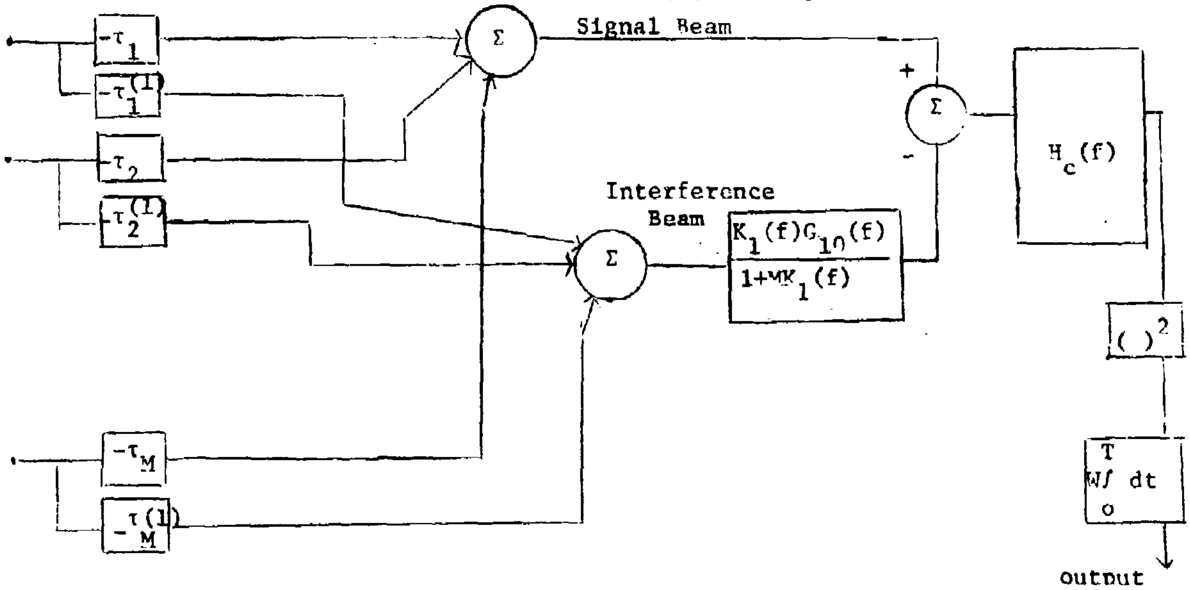


Figure 2. Optimum Filter for Single Interference.

This filter can be constructed using M tapped delay lines to generate the delays τ_k and $\tau_k^{(1)}$ in each hydrophone channel, and two additional delay lines to generate the filter functions $H_c(f)$ and $K_1(f)G_{10}(f)/[1+MK_1(f)]$. The use of tapped delay lines for the construction of variable filters is discussed in detail in Refs.[2] and [3]. and it has the advantage that they permit automatic adjustment by relatively simple adaptive algorithms.

It is clear that the delay lines used in each of the hydrophone channels must have a sufficient number of taps and a sufficiently small inter-tap spacing to permit steering to any one of the distinct beams that can be resolved by the array system. In this connection it should be noted that if interference elimination were not a factor mechanical steering of the array could be used to reduce the length of these delay lines. However, since interferences may come from any direction, interference elimination requires that delay lines of the maximum length needed to steer the array through 360° be used in each channel.

A discussion of delay-line characteristics is given in Appendix B, and it is shown there that the number of taps needed tends to be very large. Specifically, for a linear array with M hydrophones spaced uniformly a distance d apart the number of taps is given by

$$K = \frac{B\ d\ M(M-1)}{2\sqrt{6}\ c}$$

where B is the signal bandwidth and c the velocity of sound. Using typical values of $B = 2\pi \times 5000$ rad/sec., $d = 2$ft, $c = 5000$ ft/sec, and $M = 100$, this works out to $K = 26400$ taps. Also, it is shown in Appendix B that the tap increment under these conditions should be on the order of 1.55 μ sec. The number of taps needed appears to be well beyond currently available hardware.

The function $G_{10}(f)$ given in Eq.(15) is most easily constructed from a delay line having at least M taps giving the delays $\tau_k^{(1)} - \tau_k$; $k = 1...M$. Since both target and interference can be located anywhere in azimuth this line must have the same resolution as that needed in the individual channels. Furthermore, the maximum delay needed is at least twice that needed in the channels. This is easily demonstrated by considering a linear array in which

$$\tau_k^{(1)} - \tau_k = k \frac{d}{c}(\sin\theta_1 - \sin \theta)$$

where θ_1 is the interference direction and θ the signal direction. The maximum value of delay, obtained for $k = M$, $\theta_1 = -\pi/2$ and $\theta = \pi/2$ is $2Md/c$ while the minimum value is $-2Md/c$. A delay line can only produce positive delays: the effect of negative delays can be obtained by inserting a fixed delay line into the line from the upper summing junction in Fig. 2, as discussed in Ref.5. If this is done the range of delays needed in the tapped delay line under consideration is $2Md/c$ compared to a maximum delay of $(M-1)d/c$ needed in the channels.

The additional frequency weighting $K_1(f)/[1 + MK_1(f)]$ is a relatively minor modification in the filter characteristic. It can be implemented by applying different weights to the tap outputs before they are summed, as explained in Refs.[2] and [3]. Thus the entire filter function $K_1(f)G_{10}(f)/[1 + MK_1(f)]$ can be constructed from a standard tapped delay line filter using a delay line with twice as many taps, hence twice as long, as the delay lines used in the channel. It therefore involves no major additional design problems.

Actually, it is possible to redraw the block diagram of Figure 2 in such a way as to cut the length of the delay line needed to generate $G_{10}(f)$ in half; i.e. to make it no longer than the lines used in the individual channels to steer the array. Such a block diagram is shown in Figure 3.

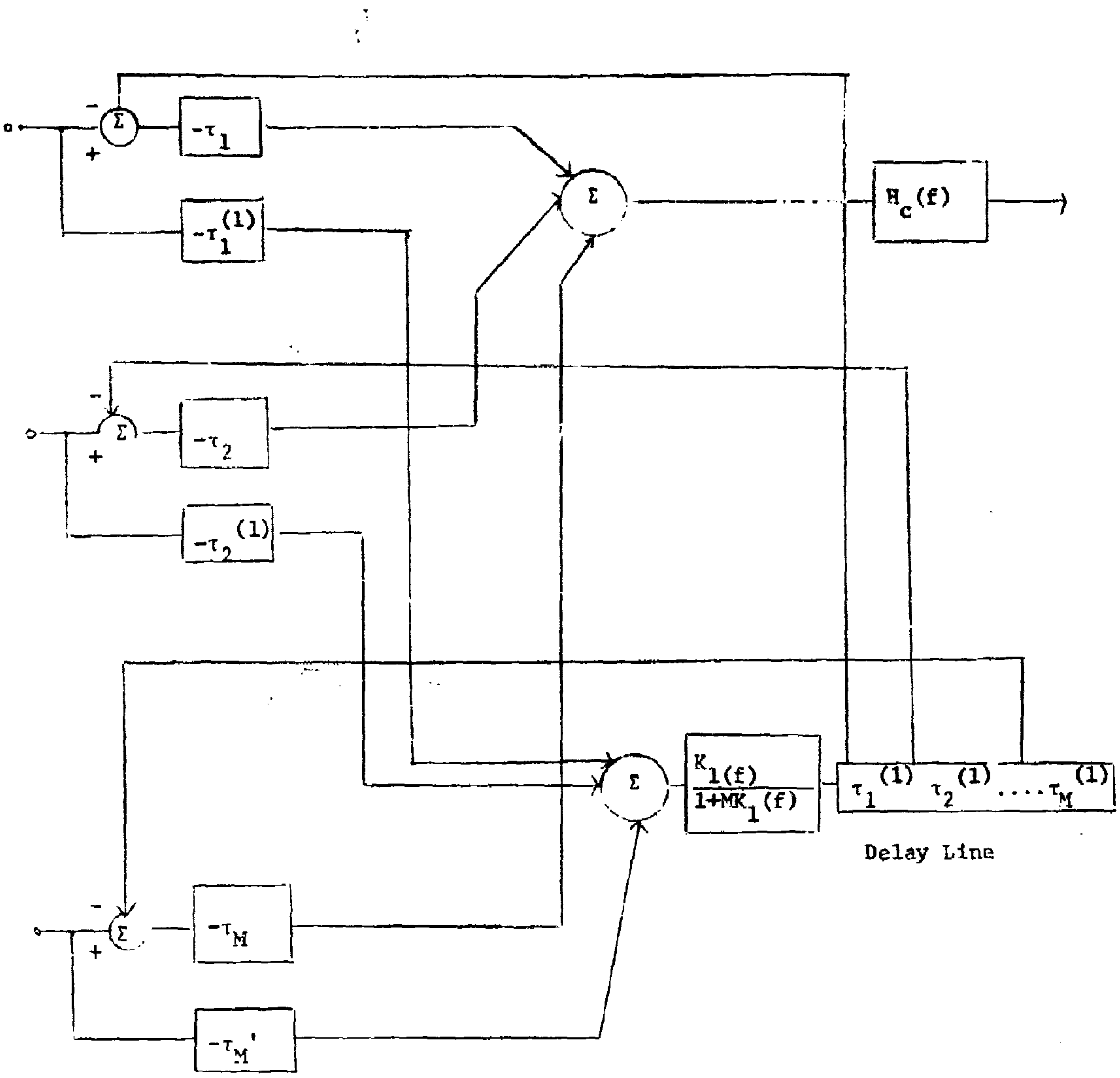


Figure 3: Modified Block Diagram for Single-Interference Filter.

It can be verified that this system produces the same output as Figure 2, but the delay line at the lower right of the diagram, which is needed to generate the function $G_{10}(f)$ has taps only at delays $\tau_1^{(1)}$, $\tau_2^{(1)}$, etc. rather than at the delay differences $\tau_1^{(1)} - \tau_1$, $\tau_2^{(1)} - \tau_2$, etc. Hence the maximum length of this line needs to be no greater than the lines used in the individual channels.

Finally, the function $H_c(f)$ must be considered. Since this is a much simpler function than $G_{10}(f)$ it can be synthesized by means of a separate tapped-delay-line filter with only a modest number of taps. Alternatively, $H_c(f)$ can be generated in each channel by summing some of the tap outputs of the delay lines used for steering the array.

If the signal and noise spectra are similar in form and if the SNR $S(f)G_o(f)/N(f)$ is small, $H_c(f)$ can be omitted entirely.

Two Interferences

With two interferences the expression in the braces of Eq.(13) becomes explicitly:

$$\underline{H}'' = \underline{Q}_o^{-1}\underline{V}_o^* - \underline{Q}_o^{-1}[\sqrt{K_1}\underline{V}_1^* \vdots \sqrt{K_2}\underline{V}_2^*] \begin{bmatrix} 1+K_1G_{11} & \sqrt{K_1K_2}G_{12} \\ \sqrt{K_1K_2}G_{12}^* & 1+K_2G_{22} \end{bmatrix}^{-1} \begin{bmatrix} \sqrt{K_1}G_{10} \\ \vdots \\ \sqrt{K_2}G_{20} \end{bmatrix} \quad (15)$$

If it is again assumed that $\underline{Q}_o = \underline{I}$, then this becomes

$$H'' = \underline{V}_o^* - A_1(f)\underline{V}_1^* - A_2(f)\underline{V}_2^* \quad (16)$$

where
$$A_1(f) = \frac{K_1(f)[1+MK_2(f)]G_{10}(f) - K_1(f)K_2(f)G_{12}(f)G_{20}(f)}{1+M[K_1(f)+K_2(f)] + [M^2-|G_{12}(f)|^2]K_1(f)K_2(f)} \quad (17)$$

and
$$A_2(f) = \frac{K_2(f)[1+MK_1(f)]G_{20}(f) - K_1(f)K_2(f)G_{21}(f)G_{10}(f)}{1+M[K_1(f)+K_2(f)] + [M^2-|G_{12}(f)|^2]K_1(f)K_2(f)} \quad (18)$$

Thus the system takes the form shown in Figure 4. If M is large, and if the interference sources are reasonably well separated it is shown in Ref.1 that $|G_{12}(f)|^2 \simeq M$ and can therefore be neglected relative to M^2. Under these conditions the denominator factors giving

$$A_1(f) \simeq \frac{K_1(f)}{1+MK_1(f)} G_{10}(f) - \frac{K_1(f)K_2(f)G_{12}(f)G_{20}(f)}{[1+MK_1(f)][1+MK_2(f)]} \quad (19)$$

and

$$A_2(f) \simeq \frac{K_2(f)}{1+MK_2(f)} G_{20}(f) - \frac{K_1(f)K_2(f)G_{21}(f)G_{10}(f)}{[1+MK_1(f)][1+MK_2(f)]} \quad (20)$$

A further simplification can be made if one assumes that $MK_1(f)$ and $MK_2(f)$ are large. If this is not true then the interference-to-ambient noise ratio is not really large enough to make interference elimination necessary or worthwhile. With this additional assumption

$$A_1(f) \simeq \frac{G_{10}(f)}{M} - \frac{G_{12}(f)G_{20}(f)}{M^2}$$

$$= \frac{1}{M}\sum_{k=1}^{M} e^{j2\pi f(\tau_k^{(1)} - \tau_k)} - \frac{1}{M^2}\sum_{k=1}^{M}\sum_{j=1}^{M} e^{j2\pi f(\tau_k^{(1)} - \tau_k^{(2)} + \tau_j^{(2)} - \tau_j)} \tag{21}$$

and

$$A_2(f) \simeq \frac{G_{20}(f)}{M} - \frac{G_{21}(f)G_{10}(f)}{M^2}$$

$$= \frac{1}{M}\sum_{k=1}^{M} e^{j2\pi f(\tau_k^{(2)} - \tau_k)} - \frac{1}{M^2}\sum_{k=1}^{M}\sum_{j=1}^{M} e^{j2\pi f(\tau_k^{(2)} - \tau_k^{(1)} + \tau_j^{(1)} - \tau_j)} \tag{22}$$

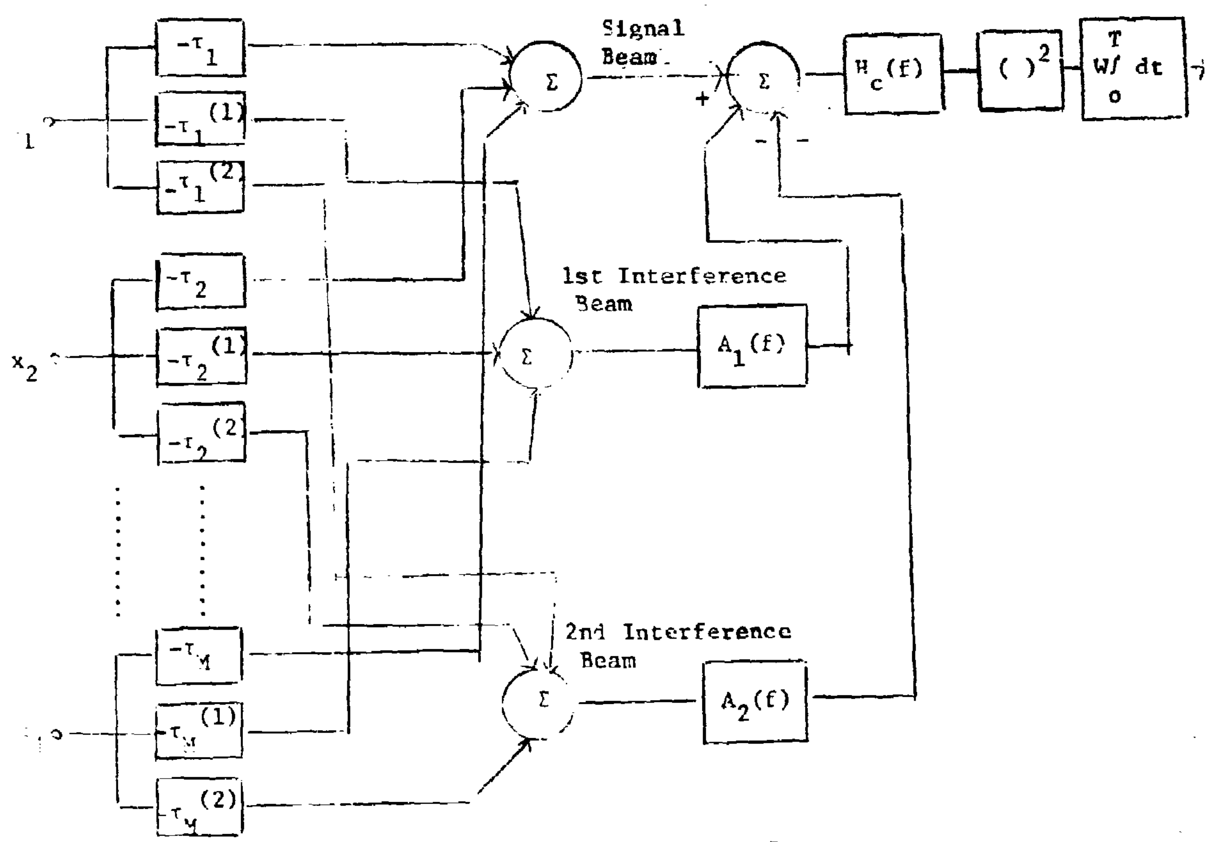


Figure 4: Filter for Two Interferences.

It is easily seen that the filter functions $A_1(f)$ and $A_2(f)$ can be constructed from tapped delay lines with weights applied to the tap outputs. Since the single summation term has already been discussed, we consider only the double summation. Again, it is clear that the resolution needed is the same as in the delay lines used in the channels. Also, if one examines all extreme values of τ_k, $\tau_k^{(1)}$, and $\tau_k^{(2)}$ one can show that for a linear array the total delay can range from -2Md/c to 2Md/c. Since the delays required by $A_1(f)$ may be the negative of those required by $A_2(f)$ it is necessary to use a fixed delay of 2Md/c in the signal beam channel and tapped delay lines of length 4Md/c in each of the interference channels. Thus the tapped delay lines must be four times as long as those used in the hydrophone channels.

Since there are M^2 terms to be summed it might appear that at least M^2 taps would be needed on these delay lines. Although it is shown in Appendix B that tapped delay lines used for linear or circular arrays should have considerably more than M^2 taps, this is not necessarily true in other array geometries. However, if a line with sufficient resolution to resolve distinct beams of the array system does not have enough taps, it simply means that some of the terms in the double summation are identical, at least to the accuracy of the delay increments. Hence these terms will be more heavily weighted in the sum. Thus it appears that the delay line required for the double summation needs to be no more complicated than that used for the single summation. Note that it is just as easy to implement the exact forms of $A_1(f)$ and $A_2(f)$ as the approximate ones given in Eqs.(21) and (22). If the additional frequency weighting required by the exact function is reasonably smooth it will call only for small changes in the weights applied to the tap outputs. This is true even if the effect of $|G_{12}(f)|^2$ in the denominator of Eqs. (17) and (18) is taken into account, because the changes introduced by this term are no more rapid than those produced by the numerator terms.

Also the function $H_c(f)$ may as well be combined with $A_1(f)$ and $A_2(f)$. Thus the optimum filter capable of handling two interference signals would consist of M tapped delay lines of unit length and two delay lines of four times this length. In addition a fixed-length delay line would be needed in the signal-beam line as discussed in the previous example. The number of taps in a unit-length tapped delay line is that given in Eq. (A-35) of Appendix B.

It is undoubtedly possible to rearrange the block diagram to make more efficient use of the delay lines as was done in the previous example. However, since such a procedure would not reduce the complexity of the delay lines by any order of magnitude, this matter is not pursued here.

<u>More than Two Interferences</u>

If the assumption $\underline{Q}_o(f) = I$ is used in Eq.(13) the expression in braces becomes:

$$\begin{aligned} \underline{H}''(f) &= \{\underline{V}_o^* - [\sqrt{K_1}\underline{V}_1^* \vdots \sqrt{K_2}\underline{V}_2^* \vdots \cdots \vdots \sqrt{K_R}\underline{V}_R^*][I + G]^{-1}\underline{g}\} \\ &= \underline{V}_o^* - A_1(f)\underline{V}_1^* - A_2(f)\underline{V}_2^* \cdots - A_R(f)V_R^* \end{aligned} \tag{23}$$

where $A_r(f) = \sqrt{K_r}(f)$ [r^{th} element of $[\underline{I} + \underline{G}]^{-1}\underline{g}$].

Since $[\underline{I} + \underline{G}]$ is now an R dimensional matrix it is clear that instead of double summations of the sort appearing in Eqs.(21) and (22) $A_r(f)$ now involves R-fold summations. A typical form of such a summation is the three-fold summation $\sum_{k=1}^{M}\sum_{j=1}^{M}\sum_{\ell=1}^{M} e^{j2\pi f(\tau_k^{(1)} - \tau_k^{(2)} + \tau_j^{(2)} - \tau_j^{(3)} - \tau_\ell)}$.

Although it is somewhat difficult to examine all terms of this sort, it is fairly clear that the maximum delay that can occur is twice the value required to steer the array through 360°. Also, it is necessary to account for the fact that these delays can be positive or negative. Thus each one of the $A_r(f)$ [or $H_c(f)$ $A_r(f)$] can be generated by a delay line of four unit lengths. The R-fold summation will, of course, require the addition of M^R terms; however as was noted in connection with the double summation, many of

these terms are identical, and therefore the variable weights applied to each tap output should permit the filter functions to be generated without the need for a larger than normal number of taps.

Assuming that a quadruple-length delay line can be constructed by connecting four single-length lines in series we see that the total number of tapped delay lines is 4R+M. In addition at least one fixed delay line is needed to permit the generation of negative delays.

It will be noted that none of the block diagrams presented so far are in the form of Figure 1. Since the block diagrams suggested in Ref. 3 are of this form, it is of some interest to consider the arrangement shown in Figure 5, which is essentially in the form of Figure 1.

By inspection of this figure the transfer function $H_k(f)$ for $k = 1 \ldots M$ is given by

$$H_k(f) = e^{-j2\pi f\tau_k} + \sum_{r=1}^{R} e^{-j2\pi f\tau_k^{(r)}} A_r(f) \qquad (24)$$

It is clear that a tapped-delay line filter that can implement $A_r(f)$ for $r = 1 \ldots R$ will have sufficient flexibility to implement $H_k(f)$. Furthermore, the post-summation filter $H_c(f)$ shown in Figs. 1 and 5 can be moved into each of the hydrophone channels, and the delay line filter that can implement $H_k(f)$ can also implement $H_k(f)H_c(f)$. Thus it appears that a delay line of four time unit length in each hydrophone channel, with adjustable weights on each tap, should suffice to generate the optimum filter function. This arrangement would therefore call for 4M unit-length lines, where unit length refers to a line capable of providing all the delays needed to steer the array through 360°. In Ref.[1] it was suggested that the number of single interferences that can be eliminated by the kind of system discussed here is on the order of $\sqrt{M}$. Therefore, since for all $M \geq 2$ $4M > 4\sqrt{M} + M$ the block diagram of Fig.5 is less efficient in the use of delay lines than that of Fig.4. However, it is again true that no order-of-magnitude difference is involved.

Conclusion

If correlation of the isotropic noise components between adjacent hydrophones is negligible then the optimum filter is shown to consist of an array system capable of forming a signal beam and R additional beams that are steered on each one of the interference sources. After passing through rather complex filters these outputs are subtracted from the signal beam-former output and the result is then passed through a post-summation filter, squared, and averaged.

Filters of considerable complexity can be synthesized automatically by use of tapped delay lines. The tap outputs are individually weighted and then summed to provide the filter output; the weighting can be accomplished by a simple computer which implements an adptive algorithm. Such adaptive filters have been considered by Luckey [2] and by Chang and Tuteur [3].

If tapped delay lines are used to generate the R + 1 beams that must be formed by the system it is shown that the number of taps required is proportional to BD^2/cd where B is the signal bandwidth, D is the array size, d is the interphone spacing and c is the velocity of sound. Typically, for a linear array with 100 hydrophones the number of taps required is on the order of 20000 or more; which appears to be well beyond current technology. However, since this requirement arises primarily from the need to produce several beams it is shared by suboptimum processors such as the simple multiple beam former. In fact, it is also shown that the additional complication needed to make a conventional system into an optimum one is relatively minor. For a system having M hydrophones, and capable of eliminating R interference sources, the optimum system would require 4R+M delay lines, while the simple beam former would require M delay lines. The

conclusion seems to be therefore that if tapped delay lines can be built to steer the array satisfactorily then the optimum processor can be built fairly easily by the use of a few additional delay lines plus some relatively simple associated circuitry.

List of References

1. F.B. Tuteur, "The Effect of Noise Anisotropy on Detectability in an Optimum Array Processor", General Dynamics/Electric Boat Research Progress Report No. 33 (September 1967).

2. R.W. Lucky, "Automatic Equalization for Digital Communication" Bell System Technical Journal XLIV, No. 4, (April 1965), pp.547-588.

3. J.H. Chang and F.B. Tuteur, "Methods of Stochastic Approximation Applied to the Analysis of Adaptive Tapped Delay Line Filters", General Dynamics/Electric Boat Research Rept., No. 34 (October 1967).

4. V.C. Anderson, "Steerable Null Processing", Proc. 23 Naval Symposium on Underwater Acoustics 1965 (429-433).

5. H. Cox, "Array Processing Against Interference", Naval Ship Systems Command, Washington D.C., October 1967.

6. P.M. Woodward, "Probability and Information Theory with Applications to Radar", McGraw-Hill Book Company, Inc., New York, 1953, page 101.

Appendix A

The Number of Distinct Beams Produced by an Array

Consider a conventional array having M hydrophones as shown in Figure A1. For the sake of simplicity it is assumed that the only processing done is to

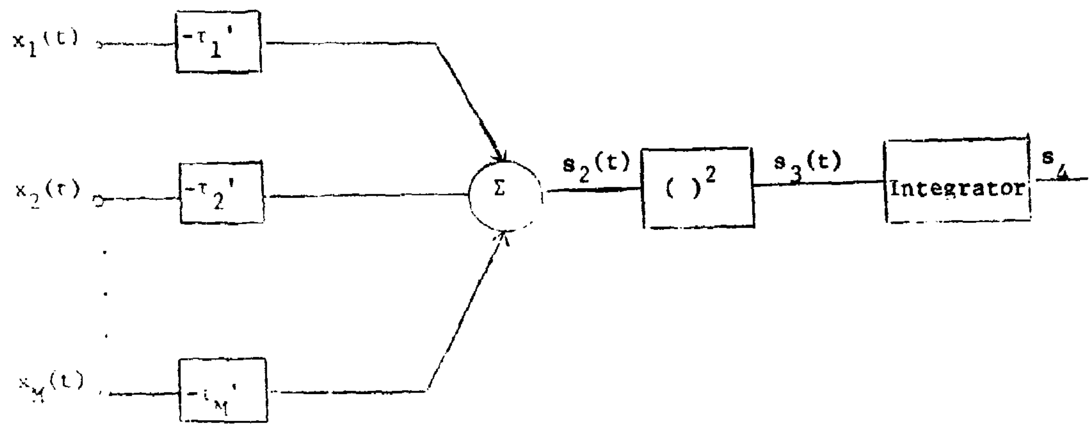


Figure A1· Conventional Array.

delay the signal from each hydrophone in order to steer the array· the delayed signals are then summed, the result is squared, and averaged.

The received signal is

$$\underline{x}(t) = \underline{s}(t) + \underline{n}(t) \qquad \text{A.1}$$

where
$$\underline{x}(t) = [x_1(t) \ldots\ldots x_M(t)]^T,$$
$$\underline{n}(t) = [n_1(t) \ldots\ldots n_M(t)]^T,$$
$$\underline{s}(t) = [s_1(t) \ldots\ldots s_M(t)]^T.$$

For the purpose of this discussion we consider only $\underline{s}(t)$, which is assumed to be expanded in a Fourier series so that

$$\underline{s}(t) = \sum_{n=-WT}^{WT} s_n \underline{V}(n)\, e^{j\omega_n t} \qquad \text{A.2}$$

where $\underline{V}(n) = \begin{bmatrix} e^{j\omega_n \tau_1} \\ \vdots \\ e^{j\omega_n \tau_M} \end{bmatrix}$ gives the signal direction.

The effect of the delays following the hydrophones is to multiply the signal vector by a steering matrix with the result that after summation the signal $s_2(t)$ is given by

$$s_2(t) = \sum_{k=1}^{M} \sum_{n=-Wt}^{WT} s_n e^{j\omega_n(\tau_k - \tau_k')} e^{j\omega_n t} \qquad \text{A.3}$$

where τ_k' is the delay imparted to the signal at the k^{th} hydrophone. After squaring

$$s_3(t) = \sum_{k=1}^{M} \sum_{\ell=1}^{M} \sum_{n=-TW}^{TW} \sum_{m=-TW}^{TW} s_n s_m e^{j[\omega_n(t+\tau_k-\tau_k')+\omega_m(t+\tau_\ell-\tau_\ell')]} \qquad \text{A.4}$$

The effect of the integrator is to pass only the dc component of this signal which is obtained by setting $m = -n$ in the last summation (we make use of the usual convention that $\omega_{-m}=-\omega_m$). Thus

$$s_4 = \sum_{k=1}^{M} \sum_{\ell=1}^{M} \sum_{n=-TW}^{TW} |s_n|^2 e^{j\omega_n[\tau_k-\tau_k'-\tau_\ell+\tau_\ell']} \qquad \text{A.5}$$

$$= 2 \sum_{k=1}^{M} \sum_{\ell=1}^{M} \sum_{n=1}^{M} |s_n|^2 \cos \omega_n(\tau_k - \tau_k' - \tau_\ell' + \tau_\ell) \qquad \text{A.6}$$

where in the last summand the term corresponding to $n = 0$ has been omitted. The summation over n can be approximated by an integral in which $|s_n|^2$ becomes $S(\omega)$, the signal power spectrum; thus

$$s_4 = \frac{T}{\pi} \sum_{k=1}^{M} \sum_{\ell=1}^{M} \int_0^{2\pi W} S(\omega) \cos \omega(\tau_k-\tau_k'-\tau_\ell+\tau_\ell')d\omega \qquad \text{A.7}$$

The argument of the cosine function is a function of the steering angle θ which depends on the array geometry. Typically, for a linear array in which the hydrophone spacing is d

$$\tau_k - \tau_k' - \tau_\ell + \tau_\ell' = \frac{d}{c} (k-\ell)(\sin \theta - \sin \theta') \qquad \text{A.8}$$

where c is the velocity of sound, θ is the direction of the signal, and θ' is the direction in which the array is steered. For other array geometries one can in general only say that

$$\tau_k - \tau_k' - \tau_\ell - \tau_\ell' = \frac{D}{c} f(k,\ell,\theta,\theta') \qquad \text{A.9}$$

where D is some distance parameter (such as the diameter in a spherical array) and where $f(k,\ell,\theta,\theta')$ is a dimensionless function having the following properties

$$\begin{aligned} f(k,\ell,\theta,\theta) &= 0 \qquad \text{for all } k \text{ and } \ell \\ f(k,k,\theta,\theta') &= 0 \qquad \text{for all } \theta \text{ and } \theta' \end{aligned} \qquad \text{A.10}$$

If the array is steered approximately in the signal direction $\theta' = \theta + \Delta\theta$, where $\Delta\theta$ is small, then

$$f(k,\ell,\theta,\theta') = \Delta\theta\, f'(k,\ell,\theta) + \frac{(\Delta\theta)^2}{2} f''(k,\ell,\theta) + \ldots \qquad \text{A.11}$$

where $f'(k,\ell,\theta) = \frac{d}{d(\Delta\theta)} f(k,\ell,\theta,\ \theta + \Delta\theta)\Big|_{\Delta\theta=0}$ etc.

If $\Delta\theta$ is small, and if $f'(k,\ell,\theta)$ is finite, only the first term of this series needs to be retained, so that for small $\Delta\theta$ Eq.(A.7) becomes:

$$s_4(\Delta\theta) = \frac{T}{\pi} \sum_{k=1}^{\ell} \sum_{\ell=1}^{\ell} \int_0^{2\pi W} S(\omega)\cos[\omega\frac{D}{c} f'(k,\ell,\theta)\Delta\theta]d\omega \qquad \text{A.12}$$

It is now also possible to expand s_4 in terms of $\Delta\theta$ around $\Delta\theta = 0$. If only terms up to second order are retained in this expansion, then in view of A.10

$$s_4(\Delta\theta) = \frac{T}{\pi} \int_0^{2\pi W} S(\omega)\{M^2 - \frac{\omega^2 D^2(\Delta\theta)^2}{2c^2} \sum_{k=1}^{M} \sum_{\ell=1}^{M} [f'(k,\ell,\theta)]^2\}d\omega \qquad \text{A.13}$$

The ratio of output for $\Delta\theta \neq 0$ to that for $\Delta\theta = 0$ is

$$\frac{s_4(\Delta\theta)}{s_4(0)} = 1 - \frac{D^2 B^2 (\Delta\theta)^2}{2c^2} \{\frac{1}{M^2} \sum_{k=1}^{M} \sum_{\ell=1}^{M} [f'(k,\ell,\theta)]^2\} \qquad \text{A.14}$$

where
$$B^2 = \frac{\int_0^{2\pi W} \omega^2 S(\omega)\, d\omega}{\int_0^{2\pi W} S(\omega)d\omega} \qquad \text{A.15}$$

When the integrals in this expression converge for $W \to \infty$ B is frequently taken as a definition of the signal bandwidth [6]. It can, of course be evaluated if an explicit form for $S(\omega)$ and a value for ω are known.

If the beam is completely off target $f(k,\ell,\theta,\theta')$ is presumably quite large and therefore the integrand in A.7 is a rapidly oscillating function for all $k \neq \ell$. Significant contributions to s_4 are therefore made only by those terms for which $k = \ell$, with the result that for the beam completely off target

$$s_4 = \frac{TM}{\pi} \int_0^{2\pi W} S(\omega)d\omega \qquad \text{A.16}$$

so that

$$\frac{s_4(\text{off target})}{s_4(0)} = \frac{1}{M} \qquad \text{A.17}$$

The beam width can now be defined in terms of the value of $\Delta\theta$ for which $s_4(\Delta\theta)/s_4(0)$ takes on some specified value between its maximum value of unity and its minimum value of $\frac{1}{M}$. We take this value to be $\frac{1}{2}$; this is a satisfactory value for all $M > 2$. For large M the double summation in A.14 is a large number and therefore the value of $\Delta\theta$ required to produce a value of $s_4(\Delta\theta)/s_4(0)$ of $\frac{1}{2}$ is small. Therefore the higher-order terms that were omitted in A.11 should, in fact be negligible for sufficiently large M.

Setting A.14 equal to $\frac{1}{2}$ results in

$$(\Delta\theta)^2 = \frac{c^2}{D^2B^2\{\frac{1}{M^2} \sum_{k=1}^{M}\sum_{\ell=1}^{M} [f'(k,\ell,\theta)]^2\}} \qquad \text{A.18}$$

The beam width is defined to be equal to $2\Delta\theta$; thus

$$\text{Beam width} = 2\Delta\theta = \frac{2cM}{DB\sqrt{\sum_{k=1}^{M}\sum_{=1}^{M} [f'(k,\ell,\theta)]^2}} \qquad \text{A.19}$$

For simple array geometries the double summation appearing in this expression can be evaluated in closed form. Thus consider a linear array in which the hydrophone spacing is d. Then letting D equal the array length, we have $D = (M - 1)d$, and we see from Eqs.A.8 and A.9 that $f(k,\ell,\theta,\theta') = \frac{1}{M-1} (k-\ell)(\sin\theta - \sin\theta')$. Hence

$$f'(k,\ell,\theta) = \frac{1}{M-1}(k-\ell)\cos\theta$$

and
$$\sum_{k=1}^{M}\sum_{\ell=1}^{M}[f'(k,\ell,\theta)]^2 = \frac{\cos^2\theta}{(M-1)^2}\sum_{k=1}^{M}\sum_{\ell=1}^{M}(k-\ell)^2 = \frac{\cos^2\theta}{(M-1)^2}\left(\frac{M^4-M^2}{6}\right)$$

Hence for the linear array

$$2\Delta\theta = \frac{2\sqrt{6}c(M-1)}{DB\sqrt{M^2-1}\cos\theta} \simeq \frac{2\sqrt{6}c}{MB\ d\cos\theta} \qquad \text{A.20}$$

if $M >> 1$.

The fact that this expression becomes infinite for $\theta = \pm\frac{\pi}{2}$ is a reflection of the fact that in the end-fire direction the first-order term in A.11 vanishes, so that the quadratic term should be used. This is a peculiarity of the linear array which does not occur with other arrays.

The number of distinct beams is most reasonably defined as

$$N = \frac{2\pi}{\text{average beam width}}$$

However, in order to avoid the complication introduced by the infinity that occurs in Eq.A.20 for $\theta = \pm\frac{\pi}{2}$ we obtain an estimate of N for the linear array by use of

$$N = 2\pi\ (\text{average reciprocal beam width})$$

The average reciprocal beam width is the average of $\frac{B(D/c)\cos\theta}{2\sqrt{6}}$ over the interval $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$, and it is equal to $\frac{BD/c}{\pi\sqrt{6}}$. Thus for a linear array with hydrophone spacing d, the number of distinct beams is approximately

$$N = \frac{2}{\sqrt{6}} BM \frac{d}{c} \qquad \text{A.21}$$

Typically, we can take d = 2 ft c = 5000 ft/sec and B = $2\pi \times 5000$ rad/sec. giving

$$N = \frac{8\pi}{\sqrt{6}} M \simeq 10M \qquad \text{A.22}$$

Another simple geometry for which A.19 can be put into a closed form is the circular array with an even number of equally spaced hydrophones.

Assume that the nominal signal wavefront is perpendicular to one of the major diameters and consider a small displacement $\Delta\theta$ away from this nominal direction. Then if D is the major diameter it can be shown that

$$f'(k,\ell,\theta) = \left|\sin \frac{\pi}{M} (k-\ell)\right| \cos \frac{\pi}{M} (k+\ell -2)$$

For sufficiently large M one can replace the double summation by a double integration:

$$\sum_{k=1}^{M} \sum_{\ell=1}^{M} \sin^2 \frac{\pi}{M} (k-\ell)\cos^2 \frac{\pi}{M} (k+\ell-2) = \frac{M^2}{\pi^2} \int_0^{\pi} \int_0^{\pi} \sin^2(x-y)\cos^2(x+y)dxdy = \frac{M^2}{4} \qquad \text{A.22a}$$

Thus Eq.A.21 becomes

$$2\Delta\theta = \frac{4c}{BD} \qquad \text{A.23}$$

This appears to be independent of M, however for constant interphone distance D is a function of M; in fact for the circular array, with M large, $D \simeq Md/\pi$. (This follows since for large M, Md is approximately the circumference of the circle). Hence Eq.A.23 becomes:

$$2\Delta\theta = \frac{4\pi}{BM^d/c} \qquad \text{A.24}$$

and the number of distinct beams is

$$N = \frac{1}{2} BM^d/c \qquad \text{A.25}$$

These expressions are very similar to the corresponding ones for the linear array, Eqs. A.20 and A.21.

The dependence of N on M is seen to be a direct consequence of the fact that both the linear and circular arrays are one-dimensional, so that for constant interphone distance D is proportional to M. This dependence is different for arrays in which the hydrophones are distributed over an area or a volume.

The simplest example of an area distribution is an array in which the hydrophones are equally distributed over a square. Such a distribution, for

M = 9 is shown in Fig. A.2.

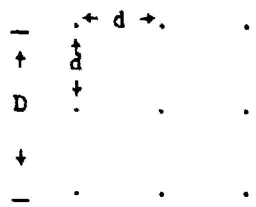


Figure A.2: Square hydrophone array.

Evaluation of $\sum_{k=1}^{M} \sum_{\ell=1}^{M} [f'(k,\ell,\theta)]^2$ is somewhat tedious, but essentially straight forward. It turns out, rather surprisingly, that the result does not depend on θ, and is, in fact exactly equal to $M^2/6$. Also, $D = (\sqrt{M}-1)d$. Hence, by use of Eq. A.19 the beam width is given by

$$2\Delta\theta = \frac{2\sqrt{6}c}{DB} = \frac{2\sqrt{6}c}{(\sqrt{M}-1)Bd} \qquad \text{A.26}$$

and therefore the number of beams is

$$N = \frac{\pi DB}{\sqrt{6}c} = \frac{\pi(\sqrt{M}-1)Bd}{\sqrt{6}c} \qquad \text{A.27}$$

Note that in terms of D,B, and c this result is essentially the same as that obtained for the linear array (Eq.A.20). Hence in going from a one-dimensional to a two dimensional array the major change is in the dependence of D on M.

From dimensional arguments this conclusion can be extended to other two dimensional arrays such as a spherical array with hydrophones only on the surface. For all such arrays

$$2\Delta\theta \propto \frac{c}{\sqrt{M}Bd} \qquad \text{A.28}$$

where the dependence on $1/\sqrt{M}$ is approximate and holds for large M. In addition for a volume distribution it is expected that

$$2\Delta\theta \propto \frac{c}{M^{1/3}Bd} \qquad \text{A.29}$$

The factor of proportionality in all of these cases appears to be on the order of 10 or less.

Appendix B

Characteristics of Delay Lines Required in Array Processors

In this appendix we examine the total time delay, number of taps, and delay between taps of the tapped-delay lines required to steer the array over 360° in azimuth. We consider initially the linear array with uniform spacing between hydrophones for the sake of mathematical simplicity; these results are then extended to other arrays with suitable modifications.

Consider a simple array processor of the form shown in Figure A.1. We assume that the signal from each hydrophones is applied to a tapped delay line and that the delays $-\tau_1'$, $-\tau_2'$,...$-\tau_M'$ are obtained by taking the output from the proper tap in each channel. We assume that the taps are equally spaced along the line, that the time delay between adjacent taps is $\Delta\tau$ and that the total number of taps is K. Thus the maximum delay that can be obtained from any line is $K\Delta\tau$. It is assumed that all the M delay lines are identical.

If a linear array is steered in the broad-side direction all the delays are equal, and we may as well assume that the delays are zero, i.e. the outputs of the first tap on each line are connected to the summer. Suppose now that we wish to steer the array away from the broad-side direction by the angle $\Delta\theta$. The smallest value of $\Delta\theta$ is obtained by making the delay of the i^{th} hydrophone equal to $(i-1)\Delta\tau$; i.e. on the first delay line we connect to the first tap, on the second delay line to the second tap, etc. For a linear array with uniform hydrophone spacing d the difference in time delay between the i^{th} and j^{th} hydrophone is given by

$$\tau_i - \tau_j = (i - j) \frac{d}{c} \sin \theta \qquad A.30$$

where c is the velocity of sound and θ is the angle between the wave front and the array axis. For small θ near θ = zero $\sin\theta \doteq \theta = \Delta\theta$. Thus, since for adjacent hydrophones the minimum value of $\tau_i - \tau_j$ is $\Delta\tau$, we have

$$\Delta\tau \doteq \frac{d}{c}\,\Delta\theta_{min} \qquad \text{A.31}$$

The minimum value of $\Delta\theta$ obtainable from a linear array with M hydrophones is given by Eq. A.20 of Appendix A. It seems reasonable to design the system in such a way that this minimum is matched to the minimum obtainable due to the limitations imposed by the finite number of taps available on the tapped delay lines. Thus we get

$$\Delta\tau = \frac{2\sqrt{6}}{BM} \qquad \text{A.32}$$

(Note that we are actually equating the $\Delta\theta_{min}$ of A.31 to $2\Delta\theta$ of Eq. A.20; however this is consistent with the definition of the number of distinct beams in Appendix A).

In order to steer the beam into the end-fire direction the delays between adjacent hydrophones must be made equal to $\frac{d}{c}$; thus the maximum amount of delay, required at the last hydrophone, is $(M-1)^d/c$. Therefore the number of taps on each delay line must be

$$K = \frac{(M-1)^d/c}{\Delta\tau} = \frac{B\,\frac{d}{c}\,M(M-1)}{2\sqrt{6}} \qquad \text{A.33}$$

Using the same typical values as in Appendix A, i.e. $B = 2\pi \times 5000$, $d = 2$, $c = 5000$, we obtain

$$\Delta\tau = \frac{1.55\times10^{-4}}{M}$$

and $K = 2.67\ M(M-1)$

If $M = 100$, $\Delta\tau = 1.55$ μsec and $K = 26400$ taps.

For other array geometries a relation such as A.31 will generally hold, except that if the spacing is not uniform d should be the smallest inter-hydrophone spacing. Then using Eq. A.19 we obtain in general

$$\Delta\tau = \frac{2dM}{BD\sqrt{\sum_{k=1}^{M}\sum_{\ell=1}^{M}[f'(k,\ell,\theta)]^2}} \qquad \text{A.34}$$

Also, since the maximum delay required is in general D/c, the number, K, of taps on the delay line must be

$$K = \frac{D/c}{\Delta\tau} = \frac{BD^2\sqrt{\sum_{k=1}^{M}\sum_{\ell=1}^{M}[f'(k,\ell,\theta)]^2}}{2cdM} \qquad \text{A.35}$$

As is shown in Appendix A the expression under the square root is generally proportional to M^2; e.g., for the circular array it is $M^2/4$. Thus it appears to be generally true that

$$K \propto \frac{BD^2}{2cd} \qquad \text{A.36}$$

with the factor of proportionality probably on the order of unity. For one-dimensional arrays, such as lines or circles, D is proportional to Md, hence K is approximately proportional to BM^2d/c. For two dimensional arrays D is proportional to $\sqrt{M}d$; therefore K is proportional to BMd/c. For volume distributed arrays, K would be proportional to $\frac{BM^{2/3}}{c}$ Similarly, the incremental delay $\Delta\tau$ is inversely proportional to BM for one-dimensional arrays to $BM^{1/2}$ for two dimensional arrays and to $BM^{1/3}$ for three-dimensional arrays.

ADAPTIVE ARRAY PROCESSORS

by

John H. Chang
Franz B. Tuteur

Progress Report No. 39

General Dynamic/Electric Boat Research

April 1969

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY

SUMMARY

This investigation is concerned with the design and analysis of an adaptive array processor in which the individual filters consist of tapped-delay lines and adjustable gains. Convergence properties of the iterative procedures are considered and the performances in filtering as well as in detection are determined analytically.

Chapter I presents the background and description of the problem to be considered. Chapter II describes the structure of tapped-delay-line filters in an array. The effect of misadjustment and the relationship between mean squared error and the number of delay elements are discussed.

In Chapter III the design of adaptive tapped-delay-line filters is formulated. The method of stochastic approximation and mean-squared-error criterion are employed to adjust the gains automatically. It is shown that it is not necessary that the desired signal generally used to obtain the error function be available. Either signal or noise correlation functions will suffice to generate the error gradient. Problems basic to all adaptive processes such as the conditions for convergence, rate of convergence, choice of the weighting sequence are answered with explicit expressions. Adaptation in a nonstationary environment is considered in Chapter IV using algorithms derived from the Kalman filtering techniques and dynamic stochastic approximation methods.

In Chapter V one approach to the design of an optimum adaptive array detection system is considered. Use is made of the convergence properties of adaptive tapped-delay-line filters and the properties of likelihood-ratio detectors for the case of Gaussian input processes and low input signal-to-noise ratios. This approach is especially useful when the received waveforms are disturbed by strong but unknown noise sources. The performances of the proposed adaptive detector are analyzed for bandlimited processes. The output signal-to noise ratio and directivity patterns are evaluated and compared with those of the nonadaptive systems.

In Chapter VI results obtained from digital computer simulations are presented to check the afore-mentioned analyses using both actual sonar signals and data generated from random numbers.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER ONE INTRODUCTION

1.1 The General Problem

The problem of designing a linear device to eliminate noise or to predict the future behavior of an incoming signal was considered by Wiener [1] more than twenty-five years ago. Wiener filters are optimal in the least square sense for stationary signals. More recent work by Kalman and Bucy [2] has led to the design of optimal time-variable linear filters for certain kind of non-stationary signals. For such signals, Kalman-Bucy filters can deliver substantially better performance than Wiener filters.

Both the Wiener and the Kalman-Bucy filters must be designed on the basis of a priori or assumed knowledge about the statistics of the input (useful signals and noises) to be processed. These filters are optimum in practice only when the statistical properties of the actual input signals match the a priori information on the basis of which the filters were designed. When the a priori information is not known perfectly, the filters will not deliver optimal performance. The concept of adaptive filters has been developed to solve such problems.

An adaptive filter can adapt itself to changing operating conditions. These changes may be due to variations in the input signals or the internal structure of the filter. Adaptation is accomplished by obsevation of the reaction of the filter to an external signal or to an internal variation with subsequent goal-directed variation of the filter parameters so that some quality criterion is minimized.

There are several criteria for optimization of a processor for an array of sensors such as hydrophones. Farran and Hills [3] have used the criterion of maximization of array gains to design real weightings for individual sensors. With a similar approach Mermoz [4] has been concerned with the optimum utilization of an array for separation of a signal of known waveform from noise. Wiener [1] used the criterion of minimizing signal distortion to design filters. Burg et al [5] developed a theory for spatial processing of seismometer arrays based on the Wiener least-squared-error criterion. Bryn [6] used the evaluation of the Neyman-Pearson likelihood ratio to minimize risk, whereby a theory of optimum signal processing has been developed for three-dimensional arrays operating on Gaussian signals and noises. All of the above contributors were concerned with matrix-inversion techniques for the optimum solution to the array processing problems. Edelblute, Fisk, and Kinneson [7] have shown that the above criteria yield equivalent results at a single frequency. Performance comparison between optimum, suboptimum, and conventional detection systems under different operating situations has been made recently by Schultheiss and Tuteur [8,9,10].

When the noise or signal distribution is not perfectly known, the afore-mentioned detection methods present two major difficulties. If the underlying statistics are unknown, the previous techniques cannot be used, if they are incorrectly assumed, the consequent detector performance can be absurd.

Since adaptive filters can be constructed with only partial knowledge about the system and filters can be incorporated to realize most detection systems, adaptive detectors can be designed in a similar

fashion. In this study one approach to the design of an optimum adaptive array detection system is considered. Use is made of the convergence properties of adaptive tapped-delay-line filters and the properties of likelihood-ratio detectors for the cases of Gaussian input processes and low signal-to-noise ratios. This approach is especially useful when the received waveforms are disturbed by strong but unknown noise sources.

1.2 Adaptive Filters, Detectors, and State of the Art

Considerable interest has been expressed recently in the application of adaptive filters to communication problems. Widrow [11] and Gabor, et al [12] have independently investigated and constructed systems that "learn" or adjust themselves to stochastic signals in order to minimize error power. Both compare a filtered, noisy signal with a noise-free signal to obtain the error. The mean-sqare error as a function of certain of the filter parameters is a high-order parabolic surface. These parameters are adjusted according to surface searching procedures for minimum error. Gabor and Widrow each have constructed their self-organizing systems in the form of a highly specialized analog computer. Bucy and Follin [3] suggested an adaptive filter which measures the spectral densities of the imput signal and noise processes and adjusts its band-pass to give optimum filtering in the Wiener sense.

Narendra and McBride [4] described an optimization technique which is applicable to filtering problems. The change in each parameter is determined from an error gradient in parameter space computed by cross-correlation methods which are independent of signal spectra and require no test signal or parameter perturbation. This method works if either the noiseless signal is available or the signal spectrum is known. Some averaging operation is performed to obtain the parameter increments.

More recently Widrow [15,16] analyzed an adaptive filter consisting of tapped-delay-line and adjustable gains. The adaptive algorthm was obtained through heuristic reasoning rather than mathematical rigorousness. Some approximate methods were given to estimate the rate of adaptation, effect of misadjustments, etc. However, a noise-free signal or a simulated signal is required to adjust the gains.

A number of authors have applied "adaptive" techniques to the problems of detecting signals in the presence of noise. The problem of designing an adaptive filter for a fixed waveform whose time arrival is unknown has been considered by Glaser [18]. In his work a statistical decision theory approach is used. Local waveform uncertainty is expressed in terms of an a priori probability density function but recurrence time uncertainty is not. The epoch is instead detected on a local basis and the assumption is made that epoch measurement is accurate.

Jakowatz, Shuey, and White [19] have proposed an adaptive filter for detecting a recurrent fixed waveform. The basic operations are : (1) comparison of a sample of the incoming waveform with an estimate of the unknown signal, (2) correlation of these two, (3) on the basis of the correlator output, guess whether or not a signal is contained in the current sample of the incoming waveform, (4) at those times when a signal is guessed to be present, form a new estimate of the signal which consists of a weighted average of that sample of the input with the prior estimate.

Although basic guidelines from signal detection theory are used in the adaptive filter of Jakowatz and et al, the design approach is not an optimum one as the authors indeed recognized. Two characteristic features are apparent in this adaptive filter. First, a local detection is required before any modification of the memory is made. Secondly, the receiver memory is used to remember a single waveform. This is undoubtedly an inadequate memory for the receiver to be optimum. Their adaptive filter may be, however, a practical receiver when the local waveform signal-to-noise ratio is large enough to permit good local

detection. In such a case the simple implementation of a receiver with a single waveform memory may justify its suboptimum detection performance.

Daly [20] and Scudder [21] have considered a local detection problem in which a fixed local waveform recurs in a synchronous manner. In the local detection case the problem becomes that of detection where each of the local waveform recurrences are using all the past information. The approach is Bayesian and one of optimum receiver design. One central problem is common, however, and that is the problem of implementing an optimum receiver which requires an exponentially growing memory. As Scudder [21] pointed out, the standard nonsequential realization of the optimum receiver is very complex, grows exponentially with time, and the analysis of its performance is close to impossible, even using present day computers.

In detection problems, the primary goal is to decide between two hypotheses: presence of signal plus noise or noise alone. If one prefers correct decision to mistakes, Peterson, Birdsall, and Fox [22] have shown that the optimum receiver is one which realizes the likelihood ratio of the observation and this fact does not depend on any specific quantity to be maximized or minimized.

The likelihood ratio plays a central role in the design of adaptive receiver realization as it did in the design of optimum receivers in classical detection theory. The adaptive receiver realization in this report is obtained by constructing Wiener filters for each sensor output, cascading the sum of these filters with the inverse square root of the signal spectrum density, then squaring and averaging. Since the Wiener filters are approximated

by tapped-delay lines and realized adaptively, the detector implementation is very simple and practical. For Gaussian processes and low signal-to-noise cases, the proposed system will asymptotically form a likelihood ratio detector and at the same time the output signal-to-noise ratio is maximized.

1.3 Problem statement and objectives

The problem considered in this study is the passive detection of a noise-like signal waveform generated by a source located in a known direction from the receiver. Typical applications of this general problem can be found in sonar detection, seismic detection and radio communications. The sonal application is the one that primarily motivates this study, and examples will be taken from the sonar area. In order to take advantage of the known directivity of the target signal, a directional receiver in the form of an array is employed to distinguish signal from noise. In the sonar application, the receiver consists of an array of hydrophones, together with an appropriate processor. Generally speaking, the processor consists of individual filters on each sensor output, a summer, a post-summation filter, a square-law device and an averaging filter. The output of the averaging filter is used to indicate the presence of a target signal.

In the absence of a target signal the averaging filter output is the result of noise waveforms picked up by the array elements. The noise is partly far-field noise and partly locally generated. The far-field noise is often assumed to be directionally isotropic; however, there may also be directional noise sources. These directional noise sources are referred to as interference sources; while the directionally isotropic component is referred to as ambient

noise. In the absence of interference noise, detection of a target signal can be based simply on the presence of a directional component in the received signal. However, if interference sources can be expected to be present in the noise field, then it is necessary to define the target in some way to distinguish it from the directional noise components.

The research described herein is concerned with developing a system for processing the outputs of a passive array of hydrophones under the following assumptions:

1) Target, interferences, ambient noise and local noise are assumed to be gaussian random processes.
2) The sum of interferences and ambient noise are regarded as the effective noise, which is assumed to be statistically independent of the target signal.
3) The target-signal component $s_i(t)$ observed at the output of the i^{th} hydrophone is a linear time-invariant transformation of d(t), the target-signal component observed at the output of an ideal isotropic hydrophone located at the origin of the coordinates. The target direction is known, together with its autocorrelation function (but not necessarily its power level).
4) The statistics of the noise field are completely unknown. Interferences may be present, but this is not known. If they are present, their directions are unknown.
5) The wavefronts of target and interferences are regarded as plane over the dimensions of the receiving array.
6) The processor is a directional array whose gain is maximized in the direction from which the target is expected to come.

Since the processor is to be designed in such a way that it can be easily implemented and be able to operate well in real time in the presence of unknown noise field, adaptive techniques must be employed. The system proposed here consists of an adaptive linear multichannel filter and algorithms for iterative adjustment of the filter coefficients on the tapped-delay lines. A new philosophy is introduced here for designing adaptive algorithms using the methods of stochastic approximation. This philosophy allows any given partial information- e.g., the correlation functions between the wavefront and various delayed signals - to be incorporated directly into the weight-adjustment procedure.

This information is completely specified once the spectrum and the direction of the target are known. Since this term appears in the adjustment formula, a space-time filter optimum in a predetermined direction is produced. This filter is supposed to reduce disturbances coming from other directions. When a signal appears in the expected direction, a maximum response will show on a display device. The average bearing response can be obtained from a plot of the averaged squarer output versus the looking angle of the array. In most practical situations a narrow peak is considered to be the target.

Convergence properties of these algorithms are investigated both analytically and using simulation experiments as examples. The variations of error variance, signal-to-noise ratio, and directivity patterns during and after the adaptation period are determined.

CHAPTER TWO

GENERAL FORM OF THE ADAPTIVE PROCESSOR

2.1 Signal and Noise Models

Let us consider an array of K omni-directional hydrophones. If both target signal and noise are present at each hydrophone, the total signal received by the ith hydrophone is

$$x_i(t) = s_i(t) + n_i(t) \tag{2.1-1}$$

where $s_i(t)$ is the signal component and $n_i(t)$ is the noise component. It is assumed in all cases that the signal originates from a source sufficiently remote from the hydrophone array so that the wavefront is essentially plane over the dimensions of the array. This assumption also neglects distortions due to surface scattering and other propagation effects. Let d(t) be the signal received by a hypothetical hydrophone situated at an arbitrary reference point in the array. d(t) is assumed to be a member function of a zero-mean gaussian random process. If the array and its housing were acoustically transparent, the signal component at the ith hydrophone is $s_i(t) = d(t-\tau_i)$, where τ_i represents the propagation delay between the ith hydrophone and the reference point. Then the signals at all hydrophones can be represented by the vector

$$\underline{s}(t) = [d(t-\tau_1) \quad d(t-\tau_2) --- d(t-\tau_K)] \tag{2.1-2}$$

where K is a constant denoting the number of hydrophones in an array. If this expression is Fourier transformed, there results

$$\underline{S}(\omega) = D(\omega)\ \underline{a}^{*T}(\omega)$$

where $a^T(\omega) = (e^{j\omega\tau_1} \dots e^{j\omega\tau_K})$ (2.1-3)

and where $D(\omega)$ is the Fourier transform of d(t).

Let the spectral density of the reference signal be $\phi_d(\omega)$. Then the signal field may be represented by the cross-spectral density matrix

$$\underline{\Phi}_{ss}(\omega) = \phi_d\ \underline{a}(\omega)\ \underline{a}^{*T}(\omega) \tag{2.1-4}$$

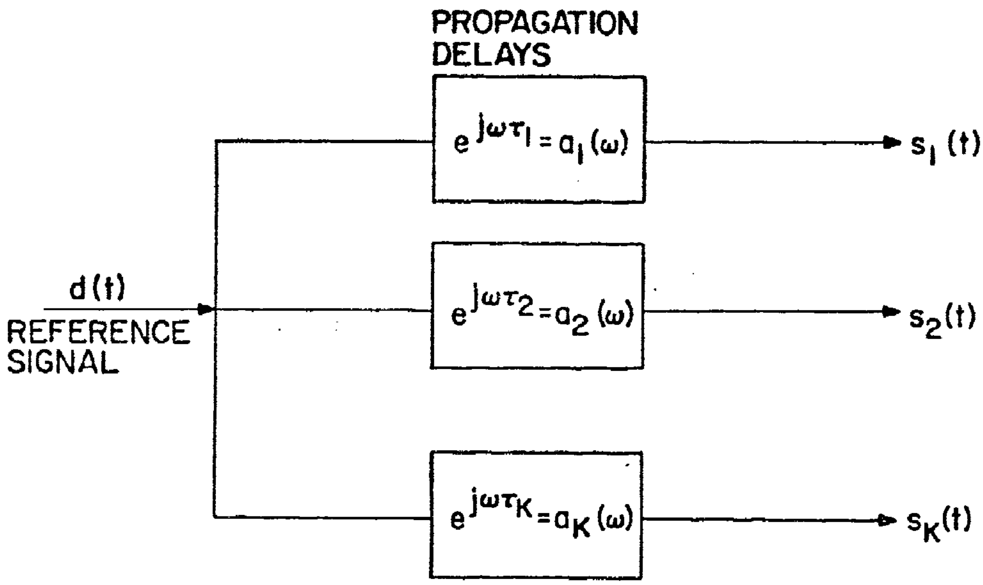


Figure 1. Signal Model

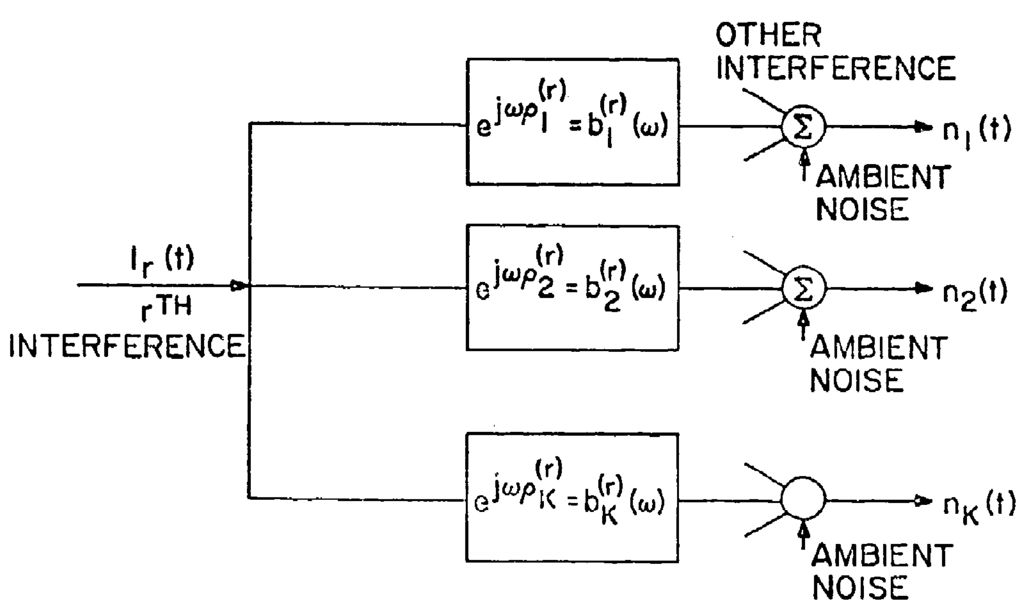


Figure 2. Noise Model

The model used for generating the observed signals is shown in Fig. 1.

The noise background is also assumed to be gaussian and to consist of ambient noise with power density matrix $\underline{\Phi}_o(\omega)$ plus interferences with power spectral density $\underline{\Phi}_i(\omega)$, $i = 1,2,\text{---}, L$. Assuming that all the interferences are statistically independent, the total noise background is then

$$\underline{\Phi}_{nn}(\omega) = \underline{\Phi}_o(\omega) + \sum_{i=1}^{L} \underline{\Phi}_i(\omega) \tag{2.1-5}$$

In case the ambient noise is independent from hydrophone to hydrophone and has power spectral density $\phi_o(\omega)$ at each hydrophone, and if there is only a single interference present with spectral density $\phi_I(\omega)$, then Eq. (2.1-5) reduces to

$$\underline{\Phi}_{nn}(\omega) = \phi_o(\omega)\ \underline{I} + \phi_I(\omega)\ \underline{b}(\omega)\ \underline{b}^{*T}(\omega) \tag{2.1-6}$$

where $\underline{I}$ is the unity matrix and

$$\underline{b}^T(\omega) = [e^{j\omega\rho_1}\ e^{j\omega\rho_2}\ \ldots\ e^{j\omega\rho_K}] \tag{2.1-7}$$

is composed of the appropriate delay for each hydrophone to steer the array conventionally at this single interference. The noise model is shown in Fig. 2.

2.2 The Structure of the Receiver

Receiving arrays consist of individual filters on each hydrophone output, a post-summation filter, a square-law device, and an averaging filter. A schematic diagram is shown in Fig. 3. They are commonly used in sonar systems to increase the ratio of desired signal power received to undesired noise power received from other sources. The hydrophones are assumed to be omnidirectional and to be passive, i.e., they receive signals from the surrounding environment. No signals are transmitted to the environment from the hydrophones.

The signal received by the array from the hydrophones are assumed to result from two separate mechanisms:

1) One component produced by the target signal propagating in the medium surrounding the hydrophones.

2) A second component produced by ambient noise and interfering sources.

The total output signal from the hydrophone to the processor is the sum of the signals described in the above if the target signal is present or just the second component in the absence of the target signal. Normally, the signal components from the individual hydrophones are related to each other through some simple linear transformation (such as a pure delay), while the noise components from these hydrophones are relatively less correlated unless some interferences are present.

The principle of the array is that, by suitably adding the outputs of individual hydrophones (perhaps after a linear transformation is applied to each), the signal components may be made to add up faster than the noise components. Then, the ratio of signal power to noise power at the summing junction or array beamformer output may be higher than at the individual hydrophone outputs. It is also true that array systems are essentially matched filters in space; a directional signal is matched by a directional receiver. The directionality of an array is obtained by properly delaying the target signal from each hydrophone and summing the result. This addition is coherent for signals coming from the direction corresponding the delays, but incoherent in other directions. Therefore, a target signal can be distinguished from the noise because of its directivity and a directional array is needed to detect it.

<u>The Optimum Receiver</u>

As shown in Fig. 3 the array processor consists of individual filter $H_i(\omega)$, $i = 1,2,---,K$, on each hydrophone, a post-summation filter $G(\omega)$, a square-law device, and an averaging filter $H_{av}(\omega)$. Although $G(\omega)$ can be included in the individual filters, it is considered separately for convenience.

There are several criteria of optimization of a processor for an array of hydrophones. It has been shown by many authors [3] - [7] and very briefly in

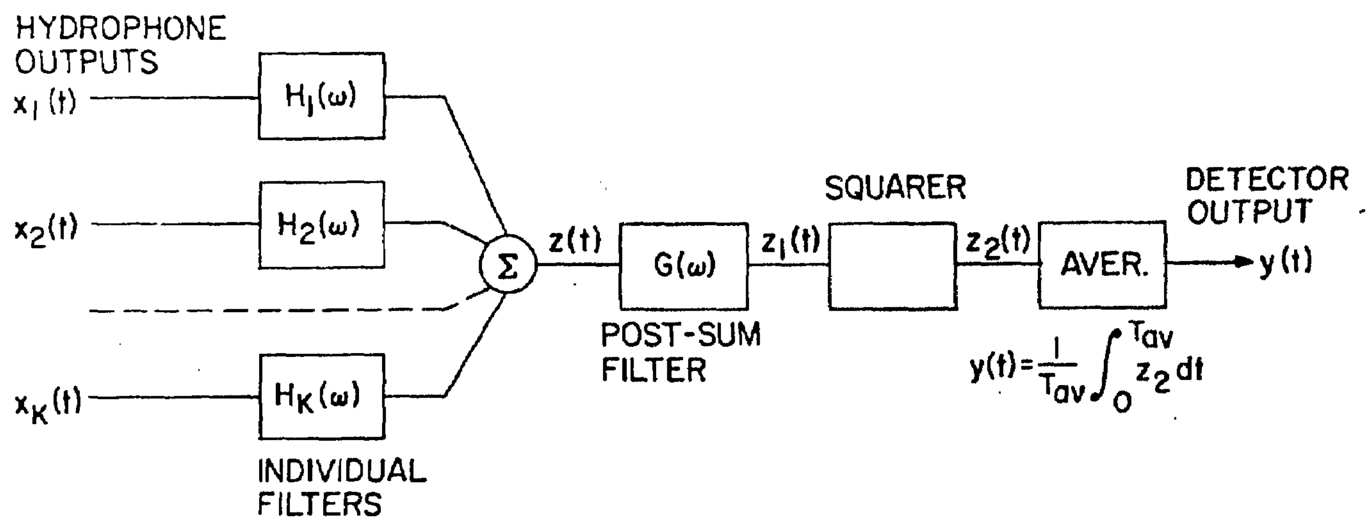


Figure 3. General Array Processor

Appendix A that the optimum individual filters

$$\underline{H}^T(\omega) = [H_1(\omega) \quad H_2(\omega) \cdots H_K(\omega)] \tag{2.2-1}$$

for the models described in Sect. 2.1 are of the form

$$\underline{H}_o(\omega) = \underline{\Phi}_{nn}^{*-1} \; \underline{a}^* \tag{2.2-2}$$

The form of the individual filters is found to be invariant under changes of optimization criteria. Only the optimum post-summation filter $G(\omega)$ needs to be modified.

Assuming Gaussian statistics for both the signal and noise and using a likelihood ratio test, $G(\omega)$ is found to be

$$G_L(\omega) = \phi_d^{\frac{1}{2}} [1 + \phi_d \, \underline{a}^{*T} \, \underline{\Phi}_{nn}^{-1} \, \underline{a}]^{-\frac{1}{2}} \tag{2.2-3}$$

If one is interested in estimation and minimizes the mean squared error between the target signal and summer output[1], the appropriate filter is

$$G_m(\omega) = G_L^2 \, (\omega) \tag{2.2-4}$$

If one maximizes the signal-to-noise ratio at the detector output $G(\omega)$ is

$$G_M(\omega) = G_m(\omega)/\phi_d^{\frac{1}{2}}(\omega) \tag{2.2-5}$$

An interesting simplification occurs for the case of small signal-to-noise ratio at the input to the squarer, or when

$$\phi_d \, \underline{a}^{*T} \, \underline{\Phi}_{nn}^{-1} \, \underline{a} << 1 \tag{2.2-6}$$

then

$$G_L(\omega) \simeq G_M(\omega) \simeq \phi_d^{\frac{1}{2}}(\omega) \tag{2.2-7}$$

and the detector structure is essentially the same regardless of the design criteria either to use a likelihood ratio test or to maximize the output signal-to-noise.

---

[1] This is really <u>not</u> what we intend to do, but Eq. (2.2-4) is included here for reference.

From Eqs. (2.2-2) and (2.2-7) we can alternatively write the optimum filters as

$$\underline{H}_{op} = \underline{\Phi}_{nn}^{*-1} \Phi_d \underline{a}^* \triangleq \underline{\Phi}_{nn}^{*-1} \underline{\Phi}_{ds}^* \tag{2.2-8}$$

and the post-summation filter

$$G(\omega) = \Phi_d^{-\frac{1}{2}}(\omega) \tag{2.2-9}$$

$\underline{\Phi}_{ds}(\omega)$, appearing in Eq. (2.2-8), is just the spectral vector between the reference signal and various signal components of the hydrophone outputs.

In any system which operates in a realistic noise field, the optimal filters must be periodically revised. It is important, therefore, that these filters assume a form that can readily be changed. The results concerning optimal filters described by Eqs. (2.2-2) and (2.2-8) have assumed that the filters are arbitrary without constraints and cannot be constructed without statistical knowledge about both the signal and the noise. In the following section we shall examine the rather practical situation in which filters in an array consist of weighted-tapped-delay lines. This type of filters can approximate the physically <u>unrealizable</u> Wiener filters such as Eq. (2.2-2) or (2.2-8) to any desirable degree. Furthermore, adaptive techniques can be applied to automatically adjust the weights on these lines without using noise statistics. The relationships between tapped-delay-line filters and the Wiener filters are discussed very shortly. The adaptive part will be treated in Chapter Three.

2.3 Tapped-Delay-Line Filters in an Array

We shall first of all describe the structure of a tapped-delay-line multichannel filter processing the outputs of K hydrophones. The output of each hydrophone enters a tapped-delay-line, and is picked off at various taps (usually equally spaced) on the delay line, delayed in time but unchanged in wave-shape. The signal from each tap is passed through an associated variable attenuator (the weight); all the attenuator output signals are then summed.

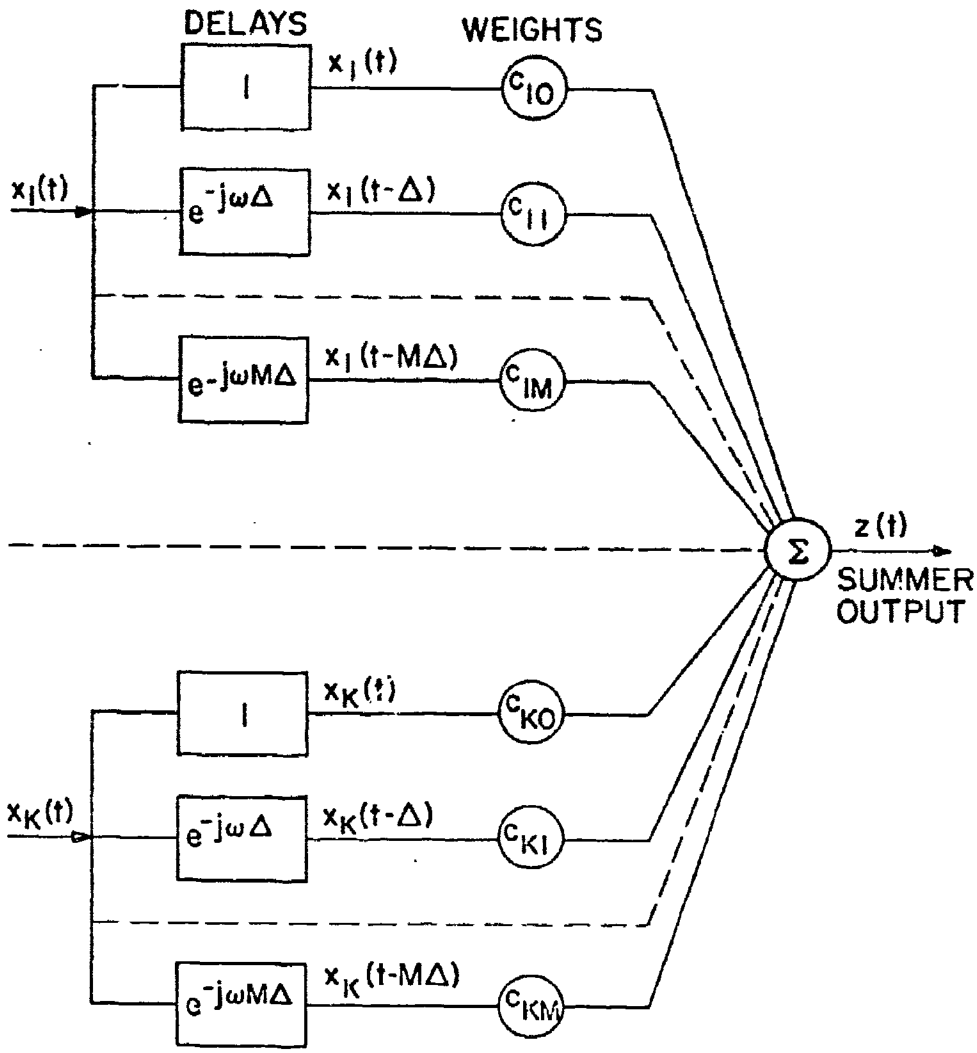


Figure 4. Tapped-Delay-Line Filter in an Array

a) Notations and the Filter Output

It is seen in Fig. 4 that each delay line consists of (M+1) tap points leading to (M+1) weights $c_{..}$ The tap points are separated by M ideal delays of Δ seconds each. Note that each weight is indexed by its tap point, and each tap point is identified by the index of the succeeding delay.

Define

$$c_{ik} = k^{th} \text{ weight on the } i^{th} \text{ filter} \tag{2.3-1}$$

$$w_{(i-1)M+k} = c_{ik} \tag{2.3-2}$$

$$\eta(t)_{(i-1)M+k} = x_i(t-k\Delta) \tag{2.3-3}$$

$$\xi(t)_{(i-1)M+k} = s_i(t-k\Delta) \tag{2.3-4}$$

$$\nu(t)_{(i-1)M+k} = n_i(t-k\Delta) \tag{2.3-5}$$

where $i = 1, 2, \ldots, K$, is the hydrophone index and $k = 0,1,2,\ldots,M$, is the tap point index. Using vector notation, the column vector of output signals $\underline{\eta}$ for the entire hydrophone array may be written as the sum of delayed target signal vector $\underline{\xi}$ and a delayed noise vector $\underline{\nu}$, or

$$\underline{\eta} = \underline{\xi} + \underline{\nu} \tag{2.3-6}$$

where

$$\underline{\eta}^T \triangleq [\eta_1(t)\ \eta_2(t) \cdots \eta(t)_{K(M+1)}] \tag{2.3-7}$$

$$\underline{\xi}^T \triangleq [\xi_1(t)\ \xi_2(t) \cdots \xi(t)_{K(M+1)}] \tag{2.3-8}$$

$$\underline{\nu} \triangleq [\nu_1(t)\ \nu_2(t) \cdots \nu(t)_{K(M+1)}] \tag{2.3-9}$$

and $\underline{\eta}^T$ denotes the transpose of $\underline{\eta}$. If the weight vector $\underline{W}$ defined as

$$\underline{W}^T = [w_1\ w_2 \cdots w_{K(M+1)}] \tag{2.3-10}$$

the filter output z(t) is then

$$z(t) = \underline{W}^T\ \underline{\eta}(t) \tag{2.3-11}$$

Note that $\underline{W}$ and $\underline{n}$ are K(M+1)-dimensional vectors.

The equations given above express the continuous time notation for the variables used in this research. Uniform discrete-time samples of these quantities are also of interest and are expressed by using the discrete-time index j. Time samples are assumed to be taken at intervals of T_{samp} seconds and, for notational simplicity, the values of the various parameters at the j^{th} sampling instant are expressed as

$$\underline{n}_j \triangleq \underline{n}(t)\Big|_{t = j\,T_{samp}} \tag{2.3-12}$$

$$z_j \triangleq z(t)\Big|_{t = j\,T_{samp}} \tag{2.3-13}$$

Because the present work is concerned with iterative weight-adjustment procedures, the j^{th} sampling instant is associated with the j^{th} iteration of the weight vector. Thus, the value of the K(M+1)-dimensional weight vector at the j^{th} iteration is $\underline{W}_j$. Hence a weight parameterized by a discrete time index j is interpreted as the j^{th} iterated value of the weight, while an unparameterized weight, as in (2.3-2), is interpreted as a time-invariant quantity.

b) Autocorrelation Matrices of the Input

When both the target signal and the noise processes are described in terms of their statistical properties, the performance of the system can be evaluated in terms of its average behavior. The quantity of most interest is the second statistical moment. For the K-dimensional vector of array output signals, $\underline{X}(t)$, the second moment becomes the (KxK)-dimensional autocorrelation matrix $\underline{R}_x(\tau)$ given by

$$\underline{R}_x(\tau) \triangleq E[\underline{X}(t)\ \underline{X}^T(t-\tau)] \tag{2.3-14}$$

where $\underline{X}^T(t) = [x_1(t)\ x_2(t)\ \cdots\ x_K(t)]$

$E[\cdot]$ denotes "expected value",

and τ is a running time-delay variable. For the K(M+1)-dimensional vector of all the signals observed at the weights, $\underline{\eta}(t)$, the second moment is the K(M+1) x K(M+1)-dimensional autocorrelation matrix $\underline{R}_\eta$ given by

$$\underline{R}_\eta(\tau) = E[\underline{\eta}(t)\ \underline{\eta}^T(t-\tau)] \tag{2.3-15}$$

Using Eq. (2.3-7) in Eq. (2.3-15) gives

$$\underline{R}_\eta(\tau) = E\left\{\begin{bmatrix}\eta_1(t)\\ \vdots\\ \eta_{K(M+1)}(t)\end{bmatrix}[\eta_1(t-\tau)\cdots\eta_{K(M+1)}(t-\tau)]\right\} \tag{2.3-16}$$

and, using Eq. (2.3-14), the second moment $\underline{R}_\eta$ becomes

$$\underline{R}_\eta(\tau) = \begin{bmatrix}\underline{R}_x(\tau) & \underline{R}_x(\tau+\Delta) & \cdots & \underline{R}_x(\tau+M\Delta)\\ \underline{R}_x(\tau-\Delta) & & & \vdots\\ \vdots & & & \\ \underline{R}_x(\tau-M\Delta) & & & \underline{R}_x(\tau)\end{bmatrix} \tag{2.3-17}$$

The above matrix is in the form of a Toeplitz matrix having equal matrix-valued elements along any diagonal. Note that by the assumption of independence of signal and noise components, we have

$$\underline{R}_\eta(\tau) = \underline{R}_\xi(\tau) + \underline{R}_\nu(\tau) \tag{2.3-18}$$

where $\underline{R}_\xi(z)$ and $\underline{R}_\nu(z)$ are, respectively, the K(M+1) x K(M+1)-dimensional signal and noise autocorrelation matrices given by

$$\underline{R}_\xi(\tau) = E[\underline{\xi}(t)\ \underline{\xi}^T(t-\tau)] \tag{2.3-19}$$

$$\underline{R}_\nu(\tau) = E[\underline{\nu}(t)\ \underline{\nu}^T(t-\tau)] \tag{2.3-20}$$

These matrices are also of the Toeplitz form analogous to Eq. (2.3-17). The advantage of the Toeplitz configuration is that the entire matrix can be constructed from the first row of the submatrices - i.e., from the matrices $\underline{R}_x(\tau)$, $\underline{R}_x(\tau+\Delta)$, $\underline{R}_x(\tau+M\Delta)$, in Eq. (2.3-14). Thus the K(M+1) x K(M+1)-dimensional auto-

correlation matrice can be stored as a $K \times K(M+1)$ -dimensional matrix:

c) Optimum Weights

The difference between the summer output and the desired (reference) signal is the error function

$$e(t) = d(t) - z(t) \tag{2.3-21}$$

using the notation defined previously we can write the square of the error as

$$e^2(t) \triangleq Q(e) = d^2 - 2\, d\underline{\eta}^T \underline{W} + \underline{W}^T \underline{\eta}\, \underline{\eta}^T \underline{W} \tag{2.3-22}$$

the mean value of which is

$$\overline{e^2} = \overline{Q(e)} = \overline{d^2} - 2\, \overline{d\underline{\eta}^T}\, \underline{W} + \underline{W}^T \underline{R}_\eta \underline{W} \tag{2.3-23}$$

To obtain the optimum vector $\underline{W}_{op}$ which minimizes the mean-squared error, we take the gradient of Eq. (2.3-23) and set the resulting form to zero,

$$\nabla_W \overline{e^2} = -2\, \overline{d\underline{\eta}} + 2\, (\underline{R}_\xi + \underline{R}_\nu)\underline{W} = 0 \tag{2.3-24}$$

or

$$\underline{W}_{op} = (\underline{R}_\xi + \underline{R}_\nu)^{-1} \underline{R}_{d\xi} \tag{2.3-25}$$

where

$$\begin{aligned} \underline{R}_{d\xi}^T &= [\overline{d(t)\eta_1(t)} \; \cdots \; \overline{d(t)\eta(t)}_{K(M+1)}] \\ &= [\overline{d(t)\xi_1(t)} \; \cdots \; \overline{d(t)\xi(t)}_{K(M+1)}] \end{aligned} \tag{2.3-26}$$

is determined completely by the signal correlation function $R_d(\tau)$ and various delays for independent signal and noises. Note that in Eq. (2.3-25) $\underline{R}_\xi$, $\underline{R}_\nu$ and $\underline{R}_{d\xi}$ are shorthand for $\underline{R}_\xi(o)$ and $\underline{R}_{d\xi}(o)$.

d) Effect of Non-Optimum Gains

The effect of non-optimum gains on the minimum mean-squared error is considered here. The absolutely minimum squared error achievable in using the

tapped-delay lines is obtained by substituting Eq. (2.3-25) into Eq. (2.3-23)

$$\overline{e^2_{min}} = \overline{d^2} - 2\, \overline{\underline{d\eta}^T}\, \underline{W}_{op} + \underline{W}_{op}^T\, \underline{R}_\eta\, \underline{W}_{op}$$

$$= \overline{d^2} - \underline{R}^T_{d\xi}\, \underline{W}_{op}$$

$$= \overline{d^2} - \underline{R}^T_{d\xi}\, \underline{R}_\eta^{-1}\, \underline{R}_{d\xi} \tag{2.3-27}$$

Using Eq. (2.3-27) and Eq. (2.3-25), Eq. (2.3-23) can be expressed as for any fixed $\underline{W}$ as

$$\overline{e^2} = \overline{d^2} - 2\, \underline{R}^T_{d\xi}\, \underline{W} + \underline{W}^T\, \underline{R}_\eta\, \underline{W}$$

$$= \overline{e^2_{min}} + \underline{R}^T_{d\xi}\, \underline{W}_{op} - 2\, \underline{R}^T_{d\xi}\, \underline{W} + \underline{W}^T\, \underline{R}_\eta\, \underline{W}$$

$$= \overline{e^2_{min}} + \underline{W}_{op}^T\, \underline{R}_\eta\, \underline{W}_{op} - 2\, \underline{W}_{op}^T\, \underline{R}_\eta\, \underline{W} + \underline{W}^T\, \underline{R}_\eta\, \underline{W}$$

$$= \overline{e^2_{min}} + (\underline{W} - \underline{W}_{op})^T\, \underline{R}_\eta (\underline{W} - \underline{W}_{op}) \tag{2.3-28}$$

If we relate the arbitrary gain to the optimum one by

$$\underline{W} = \underline{W}_{op} + \Delta\underline{W}$$

then from Eq. (2.3-28) the difference in mean squared error due to non-optimum values of $\underline{W}$ is

$$\overline{\Delta e^2} = \overline{e^2} - \overline{e^2_{min}} = (\underline{W} - \underline{W}_{op})^T\, \underline{R}_\eta\, (\underline{W} - \underline{W}_{op})$$

$$= (\Delta\underline{W})^T\, \underline{R}_\eta\, (\Delta\underline{W}) = \sum_{i=1}^{K(M+1)} \sum_{h=1}^{K(M+1)} \Delta w_i\, \Delta w_h\, \overline{\eta_i\, \eta_h}$$

$$\leq K^2 (M+1)^2 \max_{\text{all } i} |\Delta w_i| \max_{\substack{\text{all}\\ i,h}} |\overline{\eta_i\, \eta_h}| \tag{2.3-29}$$

Thus, the error due to non-optimum gains is bounded if the deviations of the gains and the input correlation functions are bounded. Note especially that for tapped-delay-line filters $\max_{\text{all } i,h} |\overline{n_i n_h}| = R_d(o) + R_n(o)$.

2.4 The Tapped-Delay Line Filters and the Wiener Filters

The multichannel Wiener filter which minimizes the mean-squared error between the summer output in an array and the target signal is obtained by combining Eqs. (2.2-2), (2.3-3) and (2.2-4). The individual filters in this case become

$$H_m(\omega) = \underline{\Phi}_{nn}^{*-1} \underline{a}^* \frac{\phi_d}{1 + \phi_d \underline{a}^{*T} \underline{\Phi}_{nn}^{-1} \underline{a}}$$

$$= [\underline{\Phi}_{ss}^* + \underline{\Phi}_{nn}^*]^{-1} \underline{a}^* \phi_d = \underline{\Phi}_{xx}^{*-1} \underline{\Phi}_{ds}^* \qquad (2.4\text{-}1)$$

where $\underline{\Phi}_{xx}(\omega)$ is the input spectral matrix

$$\underline{\Phi}_{xx}(\omega) = \underline{\Phi}_{ss}(\omega) + \underline{\Phi}_{nn}(\omega) \qquad (2.4\text{-}2)$$

and $\underline{\Phi}_{ds}(\omega)$ is the spectral density vector between the desired signal and various signal components of the hydrophone outputs.

Eq. (2.4-1) is the generalization of the single channel Wiener filter

$$H_\alpha(\omega) = \frac{\phi_d(\omega)}{\phi_d(\omega) + \phi_n(\omega)} e^{j\omega\alpha} = \frac{\phi_d(\omega)}{\phi_x(\omega)} e^{j\omega\alpha} \qquad (2.4\text{-}3)$$

for the case of long delay $\alpha \to -\infty$, [1] . The factor $e^{j\omega\alpha}$ is missing in Eq. (2.4-1), but it is understood that Wiener filters of this type are not physically realizable. Although they cannot be constructed by RLC networks, they can, however, be approximated by tapped-delay lines in practice. We will first of all show that the tapped-delay-line filters with gains given by Eq. (2.3-25) will approximate the continuous Wiener filter Eq. (2.4-1) in the

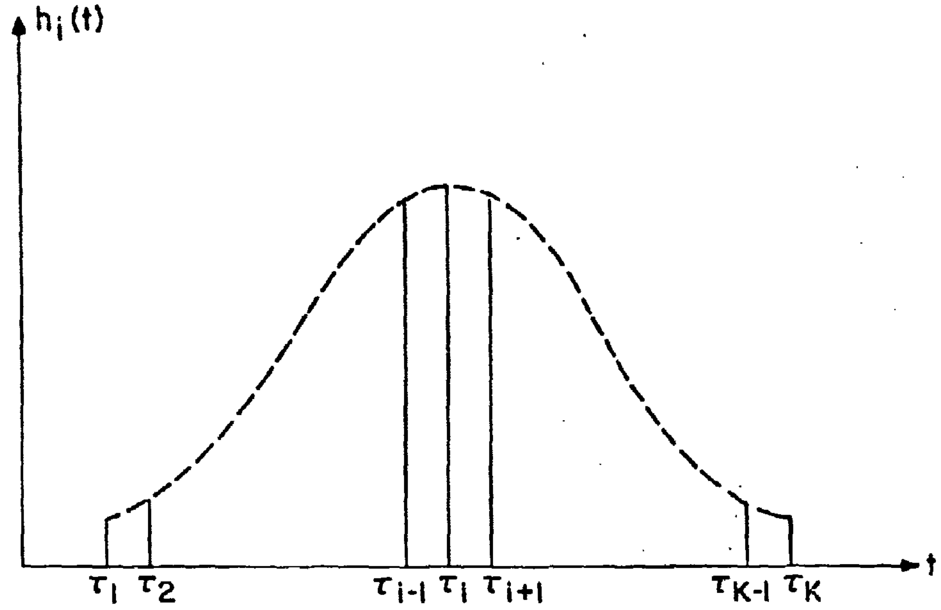


Figure 5. Impulse response of an optimum Filter (no interference)

mean squared sense.

Consider a transfer function vector

$$\underline{H}_1(\omega) = \underline{\Phi}_{xx}^{*-1}(\omega)\ \underline{\phi}_{ds}^{*}(\omega) \tag{2.4-4}$$

and the transfer function vector for the tapped-delay-line filters

$$\underline{H}_2(\omega) = \sum_{k=0}^{M} \underline{c}_k\ e^{-j\omega k\Delta} \tag{2.4-5}$$

where

$$\underline{c}_k^T = [c_{1k}\ c_{2k} \cdots c_{Kk}]\ ,\ k = 0,1,2,\cdots, M$$

Equating Eqs. (2.4-4) and (2.4-5) yields

$$\sum_{k=0}^{M} \underline{c}_k\ e^{-j\omega k\Delta} = \underline{\Phi}_{xx}^{*-1}(\omega)\ \underline{\phi}_{ds}^{*}(\omega) \tag{2.4-6}$$

Premultiplying both sides of Eq. (2.4-6) by $\underline{\Phi}_{xx}^{*}$ gives

$$\sum_{k=0}^{M} \underline{\Phi}_{xx}^{*}(\omega)\ \underline{c}_k\ e^{-j\omega k\Delta} = \underline{\phi}_{ds}^{*}(\omega) \tag{2.4-7}$$

Multiplying the above equation by $e^{j\omega\ell\Delta}$ and integrating in the frequency domain, we have

$$\sum_{k=0}^{M} \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\ \underline{\Phi}_{xx}^{*}(\omega)\ e^{j\omega(\ell-k)\Delta}\ \underline{c}_k = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\ \underline{\phi}_{ds}^{*}(\omega)\ e^{j\omega\ell\Delta} \tag{2.4-8}$$

But the frequency integrations are just correlation functions, i.e.,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\ \underline{\Phi}_{xx}^{*}(\omega)\ e^{j\omega(\ell\Delta-k\Delta)} = \begin{bmatrix} R_{x_1x_1}(\ell\Delta-k\Delta) & \cdots & R_{x_1x_K}(\ell\Delta-k\Delta) \\ \vdots & & \vdots \\ R_{x_Kx_1}(\ell\Delta-k\Delta) & \cdots & R_{x_Kx_K}(\ell\Delta-k\Delta) \end{bmatrix} = \underline{\Psi}_{xx}(\ell\Delta-k\Delta)$$

(2.4-9)

and

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\ \underline{\Phi}^{*}_{ds}(\omega)\ e^{j\omega\ell\Delta} = \begin{Bmatrix} R_{ds_1}(\ell\Delta) \\ \cdot \\ \cdot \\ \cdot \\ R_{ds_K}(\ell\Delta) \end{Bmatrix} \triangleq \underline{\Psi}_{ds}(\ell\Delta) \tag{2.4-10}$$

Thus Eq. (2.4-8) can be written as

$$\sum_{k=0}^{M} \underline{\Psi}_{xx}(\ell\Delta-k\Delta)\ \underline{c}_k = \underline{\Psi}_{ds}(\ell\Delta)\ ,\ \ell = 0,1,2,\cdots, M \tag{2.4-11}$$

or

$$\sum_{k=1}^{M}\sum_{i=1}^{K} R_{x_i x_h}(k\Delta-\ell\Delta)\ c_{ik} = R_{ds_i}(\ell\Delta) \tag{2.4-12}$$

for $\ell = 0,1,2,\cdots,M;\ h = 1,2,\cdots, K$. Using the definitions of various correlation matrices we see that Eq. (2.4-12) is equivalent to Eq. (2.3-25).

$\underline{H}_1(\omega)$ is approximated by $\underline{H}_2(\omega)$ in the sense of minimizing the mean squared error in the frequency domain. For the sake of simplicity, let us first treat the case of a single filter.

If the transfer function $H(\omega)$ can be regarded as being band limited to $(-\omega_o < \omega < \omega_o)$, then by simple Fourier expansion we have

$$H(\omega) = \sum_{k=-\infty}^{\infty} c_k\ e^{-j\omega k\Delta\omega} \tag{2.4-13}$$

where

$$\Delta = \frac{1}{2\omega_c} \tag{2.4-14}$$

and the c's are Fourier coefficients

$$c_k = \int_{-2\pi\omega_c}^{2\pi\omega_c} H(\omega)e^{jk\Delta\omega}\ \frac{d\omega}{2\pi} = \int_{-\infty}^{\infty} H(\omega)e^{jk\Delta\omega}\ \frac{d\omega}{2\pi} \tag{2.4-15}$$

The impulse response of the Wiener filter Eq. (2.4-3) generally takes the form shown in Fig. 5; i.e., it has a peak value at $\tau = \alpha$, and $|h(\tau)| \to 0$ as $|\tau - \alpha| \to \infty$. The memory of the filter can therefore be defined as the value of τ for which the ratio $|h(\tau)|/|h_{max}(\tau)|$ has some predetermined small value that is not exceeded for $\tau > \tau_o$ on the positive side and $\tau < 0$ on the negative side. If the noise spectrum is relatively flat the filter memory is proportional to the correlation time T_c of the signal.

For a filter having a finite memory $c_k \to 0$ for sufficiently large k . In general $c_{-k} = c_k^*$ and hence if $c_k \to 0$ so does c_{-k} . The infinite series of Eq. (2.4-13) can therefore be truncated to a finite series running from $k = - M/2$ to $M/2$ (where, for simplicity, M is assumed to be an even number), and the resulting finite sum will approximate $H(\omega)$ as closely as is derived by making $M+1$, the number of taps, large enough. Then

$$H(\omega) \approx \sum_{k=-M/2}^{M/2} c_k \, e^{-j\omega k\Delta} = \sum_{\ell=0}^{M} c_{\ell-M/2} \, e^{-j\omega\Delta(\ell-M/2)}$$

$$= e^{j\omega\Delta M/2} \sum_{\ell=0}^{M} c_\ell' \, e^{-j\ell\Delta\omega} \qquad (2.4\text{-}16)$$

where $\ell = k + M/2$ and $c_\ell' = c_{\ell-M/2}$.

It is readily seen that the summation terms in Eq. (2.4-16) can be constructed using weighted-tapped delay lines. The middle of the delay line corresponds to the term $k = 0$.

The minimum mean squared error resulting from the process of approximating $H(\omega)$ by $\sum_{k=M/2}^{M/2} c_k \, e^{-j\omega k\Delta}$ is obtained by choosing the coefficients in accordance with Eq. (2.4-15)

$$\overline{\varepsilon^2} = \frac{1}{2\pi}\int_{-2\pi\omega_o}^{2\pi\omega_o} \left|H(\omega) - \sum_{k=-M/2}^{M/2} c_k\, e^{-j\omega k\Delta}\right|^2 d\omega$$

$$= \frac{1}{2\pi}\int_{-2\pi\omega_o}^{2\pi\omega_o} |H(\omega)|^2 d\omega - \sum_{k=-M/2}^{M/2} |c_k|^2$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\omega)|^2\, d\omega - 2\sum_{k=0}^{M/2} |c_k|^2$$

$$= 2\sum_{k=M/2+1}^{\infty} |c_k|^2 \qquad (2.4\text{-}17)$$

Thus M is determined by the maximum possible value of $\overline{e^2}$. Since $|c_k| \to 0$ for values of k such that $k\,\Delta >> T_c$, the signal correlation time, it also follows that M is proportional to T_c/Δ with the proportionality constant chosen to produce an acceptable mean square error. A typical value of M might be $4\,T_c/\Delta$.

The extension of this argument to arrays of filters such as in Figure 3 is immediate except that in general the filters must introduce additional delays in order to steer the array. Hence the impulse response of hi(t), the ith filter, peaks at $t = \tau_i$ and diminishes for values of t away from τ_i. If all the filters consist of delay lines having M+1 taps, then by reference to Eq. 2.4-16 one can make

$$Hi(\omega) = \sum_{k=M/2}^{M/2} c_k e^{-j\omega(k+k_i)\Delta}$$

$$= e^{j\omega(M/2-k_i)\Delta} \sum_{\ell=0}^{M} c'_\ell\, e^{-j\ell\Delta\omega} \qquad (2.4\text{-}18)$$

where, as in Eq. (2.4-16), $\ell = k+M/2$, $c'_\ell = c_{\ell-M/2}$, and where in addition $k_i\Delta = \tau_i$. The value of ℓ for which c'_ℓ in maximum is then given by

$\ell = M/2 - k_i$. Also, if the maximum delay τ_i is such that $k_i \le k_{i_{max}}$ then M must clearly be increased by $k_{i_{max}}$ over the value needed for a single filter applied to the same signal spectrum, i.e., typically M might be

$$M = \frac{4T_c}{\Delta} + k_{i_{max}} \tag{2.4-19}$$

2.5 The Effect of Interferences on the Processor Structure

In this section we consider the effect of interferences on the processor structure under the assumptions that

1) The input spectra are identical in shape (but not in power levels) over the frequency range $(0,\omega_o)$ where most of the input power is concentrated.

2) The directions of the interferences are known exactly and the tap sparing are set to the desirable amount.

The optimum individual filters in a general array configuration are rewritten here for convenience

$$\underline{H}_{op} = \underline{\Phi}_{nn}^{*-1} \phi_d \underline{a}^* \tag{2.5-1}$$

where the noise spectral matrix are given by Eq. (2.1-5)

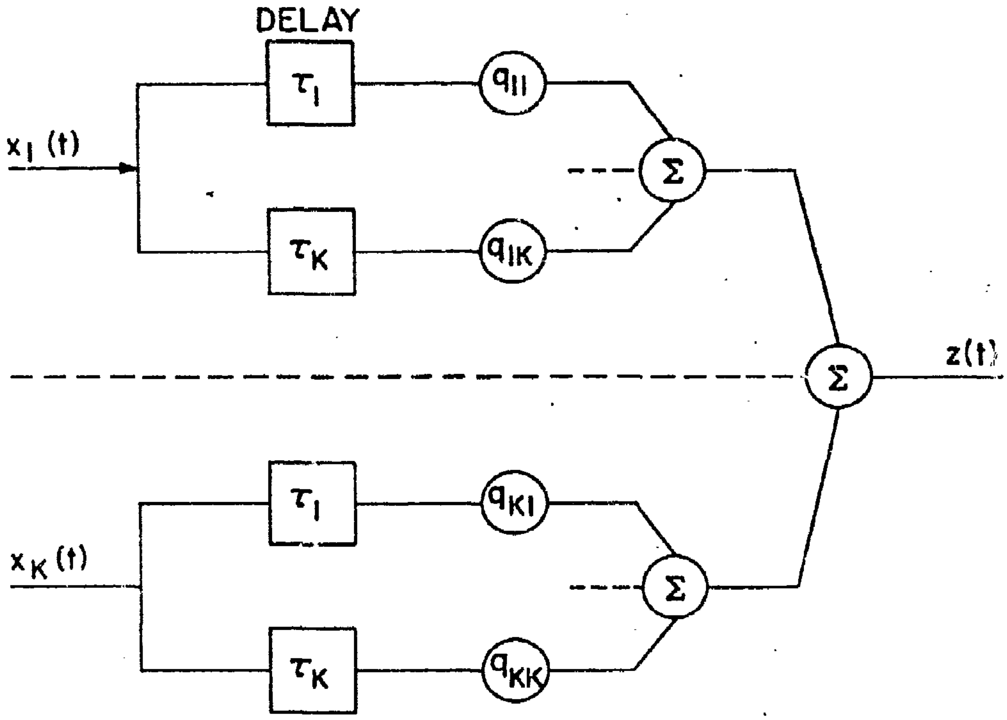
$$\underline{\Phi}_{nn} = \underline{\Phi}_c + \sum_{i=1}^{L} \underline{\Phi}_i \tag{2.5-2}$$

In the above $\underline{\Phi}$'s are understood to be function of ω ; $\underline{\Phi}_o$ is the ambient noise spectral matrix, and the $\underline{\Phi}_i (i = 1,2,\text{--}L)$ are the interference spectral matrices. The signal and noise models shown in Fig. 1 and 2 are used here. When necessary, a superscript will be used to indicate various interferences. For example, the i^{th} interference delay is defined as

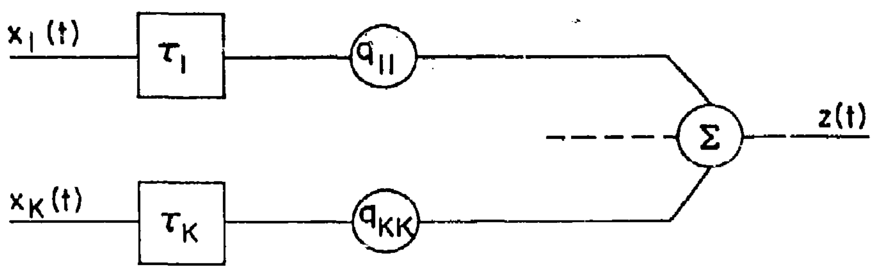
$$\underline{b}_i^T = \left\{ e^{j\omega\rho_1^{(i)}} \quad e^{j\omega\rho_2^{(i)}} \quad \cdots \quad e^{j\omega\rho_K^{(i)}} \right\},\ i = 1,2,\cdots, L .$$

a) No interference

If there is no interference, the general noise spectral matrix reduces to



(a) NO INTERFERENCE CORRELATED NOISE



(b) NO INTERFERENCE UNCORRELATED NOISE

Figure 6. Structure of Tapped-Delay-Line Filters

the ambient noise spectral matrix. Let

$$\underline{\Phi}_{nn} = \underline{\Phi}_o = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1K} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \phi_{K1} & \cdot & \cdot & \cdot\ \phi_{KK} \end{bmatrix} \tag{2.5-4}$$

The optimum filters defined in Eq. (2.2-8) become

$$\underline{H}_{op} = \phi_d\, \underline{\Phi}_o^{-1}\, \underline{a}^* = \sum_{m=1}^{K} \underline{q}_m\, e^{-j\omega\tau_m} \tag{2.5-5}$$

where

$$\underline{q}_m^T = [q_{1m}\ q_{2m} \cdots q_{Km}]$$

and

q_{ih} is the i-h element of $[\phi_d^{-1}\, \underline{\Phi}_o]^{-1}$.

Since the input spectra are assumed to be similar, the term ϕ_{ih}/ϕ_d will be a constant for $i = 1,2,\text{--},K$; $h = 1,2,\text{--},K$, K being the number of hydrophones in an array. The i^{th} optimal filter can be constructed using K taps with the weights set at

$$c_{ik} = q_{ik} \tag{2.5-6}$$

for the input signal $x_i(t)$ delayed by τ_k seconds. The implementation is shown in Fig. 6a.

If the ambient noise is independent from hydrophone to hydrophone, we then have the simpler implementation shown in Fig. 6b. Here only a single gain $c_i = \phi_d/\phi_{ii} = q_{ii}$ is used to weigh the delayed input $x_i(t-\tau_i)$. This system is similar to that studied by Schultheiss for likelihood ratio detection of Gaussian signals with noise varying from element to element of the receiving array [39]. Furthermore, if the ambient noise power is identical to all hydrophones, a conventional beamformer is obtained. Thus the cost of implementing

optimum filters depends largely on the degree of noise correlation between various hydrophones.

b) Single Interference

If there is only a single interference and the ambient noise is statistically independent from hydrophone to hydrophone, the optimum filters are

$$\underline{H}_{op} = [\phi_o \underline{I} + \phi_I \underline{b}^* \underline{b}^T]^{-1} \phi_d \underline{a}^* \tag{2.5-9}$$

and the i^{th} row is just

$$H_i(\omega) = \frac{\phi_d}{\phi_o} \left\{ e^{-j\omega\tau_i} - \frac{e^{-j\omega\rho_i}}{K+\phi_o/\phi_I} \sum_{m=1}^{K} e^{j\omega(\rho_m - \tau_m)} \right\} \tag{2.5-10}$$

If the input spectra are identical in shape, then

$$H_i(\omega) = \frac{S}{N} \left\{ e^{-j\omega\tau_i} - \frac{e^{-j\omega\rho_i}}{K + N/I} \sum_{m=1}^{M} e^{j\omega(\rho_m - \tau_m)} \right\} . \tag{2.5-11}$$

where S, N, and I are respectively the power levels of target, ambient noise, and interference. The filter defined by Eq. (2.5-11) can be constructed by setting the gains according to

$$c_{ik} = \frac{S}{N} (\delta_{ik} - \frac{1}{K + N/I}) \tag{2.5-12}$$

at taps corresponding to time delays of

$$\Delta_{ik} = \rho_i - \rho_k + \tau_k \tag{2.5-13}$$

for $i = 1,2,---, K$ and $k = 1,2,---, K$. (δ_{ik} is the Kronecker δ). Here the number of taps on each individual filter is equal to the number of hydrophones in the array ($M = K$). The effect of interference appears in summation terms of Eq. (2.5-11). The impulse response of the i^{th} filter under this situation will consist of two parts. There is a positive spike of strength $\frac{S}{N} (1 \frac{1}{K + N/I})$ at $t = \tau_i$ and negative impulses of strength $\frac{S}{N} (\frac{1}{K + N/I})$ at $\rho_i - \rho_k + \tau_k$.

c) Two or More Interferences

In the presence of L interferences, the inverse of a general noise spectral matrix has been investigated by Tuteur [33]. The results are

$$\underline{\phi}_{nn}^{-1} = \underline{\Phi}_o - \underline{\phi}_o^{-1} [\sqrt{\phi_1}\, \underline{b}_1^* \cdots \sqrt{\phi_L}\, \underline{b}_L^*][\underline{I} + \underline{G}]^{-1} \begin{pmatrix} \sqrt{\phi_1}\, \underline{b}_1^T \\ \cdot \\ \cdot \\ \sqrt{\phi_1}\, \underline{b}_L^T \end{pmatrix} \underline{\phi}_o^{-1} \qquad (2.5\text{-}14)$$

where $\underline{G}$ has elements

$$g_{ih} = \sqrt{\phi_i\, \phi_h}\; \underline{b}_i^T\, \underline{\Phi}_o^{-1}\, \underline{b}_h^* \qquad (2.5\text{-}15)$$

For independent ambient noise;

$$\underline{\Phi}_o = \phi_o\, \underline{I} \quad , \quad \underline{I} = \text{unity matrix}$$

Eq. (2.5-14) reduces to

$$\underline{\phi}_{nn}^{-1} = \underline{\phi}_o^{-1} - \frac{1}{\phi_o{}^2} [\sqrt{\phi_1}\, \underline{b}_1^* \cdots \sqrt{\phi_L}\, \underline{b}_L^*][\underline{I} + \underline{G}]^{-1} \begin{pmatrix} \sqrt{\phi_1}\, \underline{b}_1^T \\ \cdot \\ \cdot \\ \sqrt{\phi_L}\, \underline{b}_L^T \end{pmatrix} \qquad (2.5\text{-}16)$$

where

$$g_{ih} = \frac{\sqrt{\phi_i\, \phi_h}}{\phi_o}\, \underline{b}_i^T\, \underline{b}_h^* \qquad (2.5\text{-}17)$$

If two interferences are present, $L = 2$, the inverse of the corresponding noise spectral matrix becomes

$$\underline{\phi}_{nn}^{-1} = \phi_o^{-1}\, \underline{I} - \frac{\phi_o}{D} [(\phi_o + \phi_1 K)\phi_1\, \underline{b}_1^*\, \underline{b}_1^T$$

$$+ (\phi_o + \phi_2 K)\phi_2\, \underline{b}_2^*\, \underline{b}_2^T - 2\, \phi_1\, \phi_2\, (\underline{b}_1^*\, \underline{b}_2^T)^2] \qquad (2.5\text{-}18)$$

where

$$D = (\phi_o + \phi_1 K)(\phi_o + \phi_2 K) - \phi_1\, \phi_2\, |\underline{b}_1^T\, \underline{b}_2^*|^2 \qquad (2.5\text{-}19)$$

For similar input spectra, the optimal individual filters are

$$\underline{H}_{op} = \underline{\Phi}_{nn}^{*-1} \phi_d \underline{a}^*$$

$$= \frac{S}{N} \underline{a}^* - \frac{SN}{D} [(N + KI_1)I_1 \underline{b}_1^* \underline{b}_1^T \underline{a}^*$$

$$+ (N + KI_2)I_2 \underline{b}_2^* \underline{b}_2^T \underline{a}^* - 2 I_1 I_2 (\underline{b}_1^* \underline{b}_2^T)^2 \underline{a}^*] \qquad (2.5\text{-}20)$$

where I_k for $k = 1,2$ is the power of the k^{th} interference.

The i^{th} individual filter is therefore

$$\underline{H}_i = \frac{S}{N} e^{-j\omega\tau_i} - \frac{SN}{D} [(N + KI_1)I_1 e^{-j\omega\rho_i^{(1)}} \sum_{m=1}^{K} e^{j\omega(\rho_m^{(1)} - \tau_m)}$$

$$+ (N + KI_2)I_2 e^{-j\omega\rho_i^{(2)}} \sum_{m=1}^{K} e^{j\omega(\rho_m^{(2)} - \tau_m)}$$

$$- 2 I_1 I_2 e^{-j\omega\rho_i^{(1)}} \sum_{m=1}^{K} \sum_{n=1}^{K} e^{j\omega(\rho_m^{(2)} - \rho_m^{(1)} + \rho_n^{(2)} - \tau_n)}] \qquad (2.5\text{-}21)$$

In Eq. (2.5-21) the first two terms inside the bracket can be realized using K taps, but the last term would require K^2 to produce the desired impulse response. It is readily seen from Eq. (2.5-16) that in the presence of L interfering sources K^L taps would have to be used. Since the number of hydrophones in an array may be large (in the order of 10^2 or more), the cost of implementing optimal filters for several interferences could become extremely large.

A different point of view has been taken by Tuteur[45] who has considered the number of taps on the delay line required to realize the angular resolution of which the array is capable. For the particular example of a linear array he has found that the tap spacing needed to match the angular resolution is on the order of 1/BK where B is the signal bandwidth in hertz. Since the maximum delay required to steer the array over 180° in azimuth is $2(K-1)d/c$, where d is the hydrophone spacing and c the velocity of round in water, the number of taps

needed to provide all the possible signal delays required by the ρ's and τ's in Eq. (2.5-21) is

$$M = 2\,BK(K-1)d/c \tag{2.5-22}$$

Although for typical bandwidths and large array sizes this number is very large, it is independent of the number of interferences. This is true since the argument based on angular resolution implies that many of the K^L taps required to implement Eq. (2.5-16) can, in fact be considered identical. Note that M as given by Eq. (2.5-22) is on the order of $\underline{K}$ times as large as the value given in Eq. (2.4-19). The very much larger estimate obtained here is a direct result of requiring the array to be able to resolve several sources at different angles simultaneously. If delay lines with the smaller number of taps given by Eq. (2.4-19) were used one would expect a performance degradation resulting from the fact that the array could in general not be precisely steered in the various interference directions. The extent of this degradation has not been investigated.

CHAPTER THREE

THE ADAPTIVE MECHANISM

3.1 Introduction

The previous chapter presented a means for determining the optimum values of the gains provided that the statistical properties of both the desired signal and the noise are known.

In the present chapter a method is developed for adjusting the gains automatically when this information is only partially available. In particular, it will be shown that adjustment is possible if only the correlation function of the desired signal, or (not and) of the noise is available.

The adaptive filter described here bases its own design (its internal adjustable gains) upon estimated or measured statistical characteristics of input and output signals. The statistics are not measured explicitly and then used to design the filter; rather the design is accomplished in a single process by recursive algorithm which updates the adjustments with the arrival of each new data sample.

Two of the most commonly used iterative methods for making adjustments to improve system performance are the relaxation method and the method of steepest descent (or ascent). The relaxation method involves making a change in the value of only one of the controller parameters and then re-evaluating the performance measure. If the performance has been improved, a second change in the same direction is made; otherwise, the first change is retracted and a change in the opposite direction is made. This process is continued until no further improvement in the performance measure can be accomplished by adjusting that particular parameter; whereupon the same process is repeated for each of the remaining controller parameters. After several iterations through the entire procedure, the controller parameters tend toward that set of values which yields the optimum performance measure.

The methods of steepest-descent (or ascent), referred as gradient techniques are operated in a manner similar to the relaxation method, with the notable exception that all parameters are adjusted simultaneously rather than sequentially. This is done by measuring the partial derivative of the performance measure with respect to each of the controller parameters and then adjusting all the parameters in such a way that the net effect is the largest possible improvement in the performance measure. A number of techniques have been developed for determining the partial derivatives.

The most straightforward method is to perturb each of the parameters sequentially and measure the derivatives directly. This procedure, however, offers little advantage over the relaxation method. A second technique is to perturb the parameters simultaneously in such a manner that the effect of the perturbation of each parameter on the performance measure will be distinguishable from the effects of the perturbations of all the other parameters. Ways in which this may be done include perturbation by independent random noise, distinguishing the individual effects by correlation detection [14]; or perturbation by frequency-separated sinusoids, distinguishing the effects by narrow-band detection [40].

Gradient techniques can be considered as the special case of the more general method of stochastic approximation, by which either deterministic or random problems can be solved with ease. In this chapter adaptive algorithms will be derived to automatically adjust the weights on the tapped-delay lines described in the previous chapter. The methods of stochastic approximation will be used extensively in the remaining part of this research.

3.2 Methods of Stochastic Approximation

The methods of stochastic approximation were originally developed by Robbins and Monro in 1951 [28]. Their purpose was to find the root of a function disturbed by measuring noise. The term "stochastic" refers to the random character of the experimental errors, while the term "approximation" refers to the con-

tinued use of past measurements to estimate the approximate position of the goal. Kiefer and Wolfowitz [29] adapted the idea of stochastic approximation to the problem of finding the maximum of a unimodal function obscured by noise. Blum [30] used the gradient method to extend the above techniques to the multi-dimensional case. Later on Dvorestzky [31] greatly generalized and unified the whole theory and Kesten [41] derived some formulas to speed up the rate of convergence in terms of the number of changes in sign before a certain step.

a) Basic Considerations

Stochastic approximation, much like ordinary successive approximation in the absence of experimental error, involves two basic considerations--first choosing a promising direction in which to search and selecting the distance to travel in that direction. Picking a search direction is no more difficult for stochastic than for deterministic approximations, for one simply behaves as if he believed the experimental results, ignoring entirely the possibility of error. This means of course that the experimenter will move away from his goal whenever he is misled by the vagaries of chance error. It will be seen that such temporary setbacks do not prevent ultimate convergence if the step sizes are chosen properly.

In both stochastic and deterministic schemes, the corrections are made progressively small as the search proceeds so that the process will eventually converge. To make this convergence rapid, one would like to shrink the step size as speedily as possible. The main difference between stochastic and deterministic procedures is in fact the speed with which the steps can be shortened. When noise is totally absent one can reduce the steps very rapidly, but when there is danger of an occasional jump in the wrong direction, shortening the steps too rapidly could make it impossible to erase the long-run effects of a mistake. In the latter case the process would still converge, but to the wrong value.

b) The Ordinary Methods

Many problems in modern engineering systems design can be reduced to that of

finding the extrema of functions of several variables

$$I = Q(c_1, c_2, \ldots c_n) = Q(\underline{c}) \tag{3.2-1}$$

where $\underline{c} = \{c_1, c_2, \ldots, c_n\}$.

Denoting the optimal values of $\underline{c}$ by $\underline{c}_{op}$ and assuming that the extremum of interest to us is a minimum, we can obtain the solution of $\underline{c} = \underline{c}_{op}$ by setting the gradient of $Q(\underline{c})$ equal to zero, i.e.,

$$\nabla_c Q(\underline{c}) = 0 \tag{3.2-2}$$

where $\nabla_c Q(\underline{c}) = \frac{\partial Q(\underline{c})}{\partial c_1} \ldots \frac{\partial Q(\underline{c})}{\partial c_n}$.

Generally a closed-form solution cannot be obtained for (3.2-2), so iteration methods are required, especially the gradient method. The gradient method relates the coordinates of a given point with the coordinates of the preceding point and the gradient $\nabla Q(\underline{c})$. The algorithm for determining $\underline{c}_{op}$ can be written in the form

$$\underline{c}_{j+1} = \underline{c}_j - \gamma \nabla_c Q(\underline{c}_j) \tag{3.2-3}$$

When $Q(\underline{c})$ is not given analytically or is not differentiable, the gradient ∇Q can be approximately determined with the formula

$$\frac{Q_+(\underline{c}, a) - Q_-(\underline{c}, a)}{2\,a} \approx \nabla Q \tag{3.2-4}$$

where

$$Q_\pm (\underline{c}, a) = \{Q(\underline{c} \pm a\,\underline{e}_1), \ldots, Q(\underline{c} \pm a\,\underline{e}_n)\} \tag{3.2-5}$$

and $\underline{e}_i$ denotes the base vectors

$$\underline{e}_1 = \{1, 0, \ldots, 0\}\ ,\ \underline{e}_n = \{0, 0, \ldots, 1\} \tag{3.2-6}$$

The corresponding algorithm is then

$$\underline{c}_{j+1} = \underline{c}_j - \gamma \frac{Q_+(\underline{c}_j\ ,\ a_j) - Q_-(\underline{c}_j\ ,\ a_j)}{2\,a_j} \tag{3.2-7}$$

In the above we assumed that $Q(\underline{c})$ is a deterministic function. If we consider a function $Q(\underline{x}|\underline{c})$, where $\underline{x} = \{x_1, x_2, \ldots, x_n\}$ is a vector of stationary random variables with distribution $P(\underline{x})$, it is natural to attempt to find the extrema of the mathematical expectation:

$$I(\underline{c}) = \int_{\underline{x}} Q(\underline{x}|\underline{c})P(\underline{x})d\underline{x} = E_x\{(\underline{x}|\underline{c})\} \tag{3.2-8}$$

The condition for determining the optimal value $\underline{c}_{op}$ is of the form

$$\nabla I(\underline{c}) = E\{\nabla_c Q(\underline{x}|\underline{c})\} = 0 \tag{3.2-9}$$

We can apply the algorithms (3.2-3) and (3.2-7) to (3.2-9) and functional (3.2-8) only when the priori distribution $P(\underline{x})$ is known and, consequently, the mathematical expectation can be determined beforehand. Frequently, however, the probability density function $P(\underline{x})$ is unknown. Nonetheless, the optimal vector $\underline{c}_{op}$ can still be determined by applying the gradient method using $\nabla_c Q(\underline{x}|\underline{c})$ instead of $E\{\nabla_c Q(\underline{x}|\underline{c})\}$. This is one of the advantages of using the method of stochastic approximation. With this method the algorithms for determining $\underline{c}_{op}$ can be written in the form

$$\underline{c}_{j+1} = \underline{c}_j - \gamma_j \nabla_c Q(\underline{x}_j|\underline{c}_j) \tag{3.2-10}$$

if $Q(\underline{x}|\underline{c})$ is analytic and differentiable, and

$$\underline{c}_{j+1} = \underline{c}_j - \frac{\gamma_j}{2a_j} \{Q_+(\underline{x}_j|\underline{c}_j, a_j) - Q_-(\underline{x}_j|\underline{c}_j, a_j)\} \tag{3.2-11}$$

if $\nabla_c Q(\underline{x}|\underline{c})$ does not exist. Here γ_j determines the pitch of the algorithm and generally depends on the index of the step and the function itself.

Algorithm (3.2-10) is a multivariate form of the Robbins-Monro procedure, while algorithm (3.2-11) is a multivariate form of the Kiefer-Wolfowitz scheme. The analogy between deterministic and stochastic algorithms is apparent. It should be emphasized however, that stochastic algorithms deal with stationary random variables which may contain random noise in addition to the useful signal.

c) Convergence Conditions and Their Geometrical Significances

We shall describe the conditions under which the above-mentioned algorithms converge. Since the mean squared error is used throughout this study as the only performance criterion, $Q(\underline{x}|\underline{c})$ is analytic and differentiable and we therefore need to consider only algorithm (3.2-10).

Let $\underline{c}_{op}$ satisfy the equation

$$E\{\nabla_c\, Q(\underline{x}|\underline{c})\} = 0 \qquad (3.2\text{-}12)$$

$E\{\nabla_c Q(\underline{x}|\underline{c})\}$ is a set of real measurable functions of real variables c such that

$$E\{\nabla_c Q(\underline{x}|\underline{c})\} \substack{> \\ = \\ <} 0 \quad \text{for} \quad \underline{c} \substack{> \\ = \\ <} \underline{c}_{op} \qquad (3.2\text{-}13)$$

where $\underline{c}_{op}$ is a constant vector, and where $\underline{c} \gtrless \underline{c}_{op}$ means $c_i \gtrless c_{op}^{(i)}$ for all i .

Theorem 1 : Let $\gamma_1 , \gamma_2 \cdots$ be a sequence of positive numbers such that

$$\begin{aligned} &(A\ 1) \quad \lim_{j\to\infty} \gamma_j = 0 \\ &(A\ 2) \quad \sum_{j=1}^{\infty} \gamma_j = \infty \\ &(A\ 3) \quad \sum_{j=1}^{\infty} \gamma_j^2 < \infty \end{aligned} \qquad (3.2\text{-}13)$$

Let the following conditions be satisfied

$$(B) \quad \inf_{\substack{\epsilon < ||\underline{c} - \underline{c}_{op}|| < \frac{1}{\epsilon} \\ \epsilon > 0}} (\underline{c} - \underline{c}_{op})^T E\{\nabla_c\, Q(\underline{x}|\underline{c})\} > 0 \qquad (3.2\text{-}14)$$

$$(C) \quad E\{\nabla_c^T Q(\underline{x}|\underline{c})\ \nabla_c Q(\underline{x}|\underline{c})\} \le d(\underline{c}_{op}^T\, \underline{c}_{op} + \underline{c}^T \underline{c}) \qquad (3.2\text{-}15)$$

for all $\underline{c}$ in a bounded set and $d > 0$

Then the sequence c_j defined by (3.2-10) converges with probability one to $\underline{c}_{op}$.

__Proof:__ The above theorem has been proved by many authors. An outline is given in Appendix B .

We see that there are several restrictions imposed on the sequence $\{\gamma_j\}$ as well as on the behavior of the function $\nabla_c Q(\underline{x}|\underline{c})$. These conditions not only guarantee the convergence of the algorithms but also possess certain geometrical meanings.

(1) $\gamma_j > 0$. This is to assure that the corrections, on the average, are to be made in the right directions.

(2) $\gamma_j \to 0$ as $j \to \infty$. This is to assure that $\underline{c}_j$ calculated from algorithm (3.2-10) will converge on some specific value. Suppose we let the measured error gradient be $\nabla_c Q(\underline{x}|\underline{c})$ and the averaged gradient be $E\{\nabla_c Q(\underline{x}|\underline{c})\}$.
Then

$$\nabla_c Q(\underline{x}|\underline{c}_j) = E\{\nabla_c Q(\underline{x}|\underline{c}_j)\} + \underline{\xi}_j \; ; \; j = 1,2,\ldots \tag{3.2-16}$$

where $\underline{\xi}_j$ is a zero-mean random variable.

Thus, $\nabla_c Q(\underline{x}|\underline{c}_j)$ is not necessarily zero even if $\underline{c}_j = \underline{c}_{op}$. If the condition $\gamma_j \to 0$ as $j \to \infty$ is satisfied, the random fluctuation $\underline{\xi}_j$ are reduced to zero as $j \to \infty$, which permits $\underline{c}_j$ to converge.

(3) $\sum_{j=1}^{\infty} \gamma_j^2 < \infty$ or $\sum_{j=J}^{\infty} \gamma_j^2 \to 0$ as $J \to \infty$.

This condition is needed to account for the cumulative effect of the fluctuation $\underline{\xi}_j$. If Eq. (3.2-16) is substituted in Eq. (3.2-10), there results

$$\underline{c}_{j+1} - \underline{c}_j = - \gamma_j \, E\{\nabla_c Q(\underline{x}|\underline{c}_j)\} - \gamma_j \, \underline{\xi}_j \tag{3.2-17}$$

Summing the above from $j = J$ upward gives

$$\underline{c}_\infty - \underline{c}_J = - \sum_{j=J}^{\infty} \gamma_j \, E\{\nabla_c Q(\underline{x}|\underline{c}_j)\} - \sum_{j=J}^{\infty} \gamma_j \, \underline{\xi}_j \tag{3.2-18}$$

which expresses the total variation in $\underline{c}$ from the J^{th} step onward. The variance of the random part of this variation is

$$E\left\{\left(\sum_{j=J}^{\infty} \gamma_j \underline{\xi}_j\right)^2\right\} = \sum_{j=J}^{\infty} \sum_{k=J}^{\infty} \gamma_j \gamma_k E(\underline{\xi}_j \underline{\xi}_k) \tag{3.2-19}$$

It is assumed that observations on $Q(\underline{x}|c)$ are taken sufficiently far apart in time so that the $\underline{\xi}_j$ are independent.

Hence the righ-hand side of (3.2-19) becomes $\sum_{j=J}^{\infty} \gamma_j^2 \{\underline{\xi}_j^2\}$

Assuming that $E[\underline{\xi}_j^2] \leq E[\underline{\xi}_0^2]$ for all $J \leq j \leq \infty$,

$$E\left\{\left(\sum_{j=J}^{\infty} \gamma_j \underline{\xi}_j\right)^2\right\} \leq E[\underline{\xi}_0^2] \sum_{j=J}^{\infty} \gamma_j^2 \tag{3.2-20}$$

Hence the requirement $\sum_{j=J}^{\infty} \gamma_j^2 \to 0$ assures that the variance and the total random variation approach zero as $J \to \infty$.

(4) $\sum_{j=1}^{\infty} \gamma_j = \infty$.

Conditions (1) through (3) assure only that $\underline{c}_j$ converges to some value $\underline{c}_\infty$. Condition (4) assures that $\underline{c}_\infty = \underline{c}_{op}$. This follows from Eq. (3.2-18). Taking expectations on both sides yields

$$E[\underline{c}_\infty - \underline{c}_j] = - \sum_{j=J}^{\infty} \gamma_j E\{\nabla_c Q(\underline{x}|\underline{c}_j)\} \tag{3.2-21}$$

Then, since condition (4) implies $\sum_{j=J}^{\infty} \gamma_j = \infty$, if $\underline{c}_j$ approaches any value other than $\underline{c}_{op}$, $E\{\nabla_c Q(\underline{x}|\underline{c}_j)\}$ will not be zero for any $j > J$ and therefore the total corrective effort $\sum_{j=J}^{\infty} \gamma_j E\{\nabla_c Q(\underline{x}|\underline{c}_j)\}$ becomes infinite.

The above four conditions state that the rate with which γ_j decreases must be such that, on the one hand, the variance of performance index vanishes, and on the other hand, the variation in γ_j over the variation period is large enough for the law of large numbers to hold.

$$\text{(5)} \quad \inf_{\varepsilon < ||\underline{c} - \underline{c}_{op}|| < \frac{1}{\varepsilon}} (\underline{c} - \underline{c}_{op})^T E\{\nabla_c Q(\underline{x}|\underline{c})\} > 0 \quad \text{for} \quad \varepsilon > 0 \; .$$

The condition determines the behavior of the surface $E\{\nabla_c Q(\underline{x}|\underline{c})\} = y$ close to the minimum and, consequently, the sign of the increments of $\underline{c}_j$. Actually, if the error criterion does have a unique minimum, the above condition is generally satisfied.

$$\text{(6)} \quad E\{\underline{\nabla}_c^T Q(\underline{x}|\underline{c}) \; \nabla_c Q(\underline{x}|\underline{c})\} \leq d \; (\underline{c}_{op}^T \; \underline{c}_{op} + \underline{c}^T \underline{c}) \quad \text{for } d > 0 \; .$$

This condition requires that the mathematical expectation of the quadratic forms

$$E\{\nabla_c^T Q(\underline{x}|\underline{c}) \; \nabla_c Q(\underline{x}|\underline{c})\}$$

increase, as $\underline{c}$ increases, no faster than a quadratic paraboloid.

d) <u>Modification of the Ordinary Methods</u>

In this section we shall consider the algorithm

$$\underline{c}_{j+1} = \underline{c}_j - \gamma_j(\nabla Q_1 + \overline{\nabla Q}_2) \tag{3.2-22}$$

rather than the previous one

$$\underline{c}_{j+1} = \underline{c}_j - \gamma_j(\nabla Q_1 + \nabla Q_2) \tag{3.2-23}$$

where $Q_1 + Q_2 = Q$ is a function of the error, and the average of it is the performance criterion to be minimized.

Comparing Eqs. (3.2-22) and (3.2-23) with the regular gradient method with constant γ

$$\underline{c}_{j+1} = \underline{c}_j - \gamma(\overline{\nabla Q}_1 + \overline{\nabla Q}_2) \tag{3.2-24}$$

we see that in Eq. (3.2-23) no average is taken while in Eq. (3.2-22) a partial average is taken. $\underline{\xi}$, the random component of the gradient

$$\nabla Q = \overline{\nabla Q} + \underline{\xi}$$

is eliminated by the properly chosen sequence γ_j .

One would conclude therefore that the same γ_j which eliminates the error caused by the difference between $(\overline{\nabla Q}_1 + \overline{\nabla Q}_2)$ and $(\nabla Q_1 + \nabla Q_2)$ would also eliminate that caused by the difference between $(\overline{\nabla Q}_1 + \overline{\nabla Q}_2)$ and $(\nabla Q_1 + \overline{\nabla Q}_2)$. This is stated precisely in Theorem 2.

Theorem 2 : Let $\gamma_1, \gamma_2, \ldots$ be a sequence of positive numbers such that

$$\text{(A)} \quad \lim_{j\to\infty} \gamma_j = 0 \,, \quad \sum_{j=1}^{\infty} \gamma_j = \infty \,, \quad \sum_{j=1}^{\infty} \gamma_j^2 < \infty \tag{3.2-25}$$

Let the following conditions be satisfied

$$\text{(B)} \quad \inf_{\varepsilon < ||\underline{c} - \underline{c}_{op}|| < \frac{1}{\varepsilon}} (\underline{c}-\underline{c}_{op})^T E\{\nabla_c Q_1 + \overline{\nabla_c Q_2}\} > 0 \tag{3.2-26}$$

$$\text{(C)} \quad E\{(\nabla_c Q_1 + \overline{\nabla Q}_2)^T (\nabla_c Q_1 + \overline{\nabla Q}_2)\} < d(\underline{c}_{op}^T \underline{c}_{op} + \underline{c}^T \underline{c}) \tag{3.2-27}$$

where $\varepsilon > 0$, $d > 0$.

Then the algorithm

$$\underline{c}_{j+1} = \underline{c}_j - \gamma_j (\nabla Q_1 + \overline{\nabla Q}_2) \tag{3.2-28}$$

minimizing

$$E\{Q_1(e) + Q_2(e)\}$$

converges with probability one to $\underline{c}_{op}$.

Proof :

Subtracting $\underline{c}_{op}$ from both sides of Eq. (3.2-28)

$$\underline{c}_{j+1} - \underline{c}_{op} = \underline{c}_j - \underline{c}_{op} - \gamma_j (\nabla Q_1 + \overline{\nabla Q}_2) \tag{3.2-29}$$

and taking the inner product, we have

$$(\underline{c}_{j+1} - \underline{c}_{op})^T(\underline{c}_{j+1} - \underline{c}_{op}) = (\underline{c}_j - \underline{c}_{op})^T(\underline{c}_j - \underline{c}_{op})$$

$$- 2 \gamma_j(c_j - \underline{c}_{op})^T(\nabla Q_1 + \overline{\nabla Q}_2) + \gamma_j^2(\nabla Q_1 + \overline{\nabla Q}_2)^T(\nabla Q_1 + \overline{\nabla Q}_2) \tag{3.2-30}$$

Taking the conditional mathematical expectation for given $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j$ yields

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2 \; |\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j\}$$

$$= \; ||\underline{c}_j - \underline{c}_{op}||^2 - 2\,\gamma_j(\underline{c}_j - \underline{c}_{op})^T E\{\nabla Q_1 + \overline{\nabla Q}_2\}$$

$$+ \gamma_j^2 \, E\{(\nabla Q_1 + \overline{\nabla Q}_2)^T(\nabla Q_1 + \overline{\nabla Q}_2)\} \tag{3.2-31}$$

From condition (c) , Eq. (3.2-31) becomes

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2 \; |\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j\}$$

$$< ||\underline{c}_j - \underline{c}_{op}||^2 - 2\,\gamma_j(\underline{c}_j - \underline{c}_{op})^T E\{\nabla Q_1 + \overline{\nabla Q}_2\}$$

$$+ \gamma_j^2 \, d(\underline{c}_{op}{}^T \underline{c}_{op} + \underline{c}^T\underline{c}) \tag{3.2-32}$$

Using condition (B), the above reduces to

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2 \; |\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j\}$$

$$< ||\underline{c}_j - \underline{c}_{op}||^2 \,(1 + \gamma_j^2\, d) + 2\,\gamma_j\, d\, \underline{c}^T \underline{c}_{op} \tag{3.2-33}$$

From this point on, we can follow in exactly the same manner the steps leading from (B-9) to (B-18) in Appendix B.

3.3 The Design of Adaptive Tapped-Delay-Line Filters

The adaptive algorithms used here are derived from the methods of stochastic approximation stated in the previous sections. The quality criterion may be represented in the form of the mathematical expectation of some strictly convex function of the deviation of the output variation from the desired function. For simplicity we shall use the mean-squared criterion. Thus

$$I\,(\underline{c}) = E\{Q(d - z)\} \quad \text{with} \quad Q(e) = e^2 \tag{3.3-1}$$

For the tapped-delay-line filter shown schematically in Fig. 4 we know

$$x_i(t) = s_i(t) + n_i(t) \tag{3.3-2}$$

It is assumed for the moment that these functions are stationary random processes. Nonstationary or time-varying systems will be considered in a later section. As seen from Fig. 4 and the definition of d(t) given in Sect. 2.1 the error function is

$$e(t) = d(t) = d - \underline{W}^T \underline{\eta} \tag{3.3-3}$$

and its square is

$$Q(e) = e^2 = d^2 - 2\, d\underline{\eta}^T \underline{W} + \underline{W}^T \underline{\eta}\, \underline{\eta}^T \underline{W} \tag{3.3-4}$$

The gradient of Q with respect to the weights becomes

$$\nabla_W Q(e) = 2\, d\underline{\eta} + 2\, \underline{\eta}\, \underline{\eta}^T \underline{W} = -\, 2\, e\, \underline{\eta} \tag{3.3-5}$$

Upon using algorithm (3.2-10), the adaptive scheme to adjust the weights obtained as

$$\underline{W}_{j+1} = \underline{W}_j + 2\, \gamma_j\, \underline{\eta}_j\, (d_j - z_j) \tag{3.3-6}$$

The above adjustment procedure requires the availability of the error function as a real time function. This requirement is not convenient in dealing with communication problems such as filtering and detection, and it must therefore be removed.

This is done by rewriting Q(e) as

$$\begin{aligned} Q(e) &= [d(t) - z(t)]^2 = z^2(t) + d^2(t) - 2\, d(t)\, z(t) \\ &= z^2(t) + d^2(t) - 2\, d(t) \sum_{k=1}^{K(M+1)} w_k\, [\xi_k(t) + \nu_k(t)] \end{aligned} \tag{3.3-7}$$

Let

$$Q_1 = z^2(t) \tag{3.3-8}$$

$$Q_2 = d^2(t) - 2\, d(t)\, z(t) \tag{3.3-9}$$

and note

$$\nabla Q_1 = 2z\nabla z = 2\, \underline{\eta}\, z \tag{3.3-10}$$

$$E\{\nabla Q_2\} = -\,2E\{d\underline{\xi}\}\,\underline{A} - 2\,\underline{R}_{d\xi} \tag{3.3-11}$$

where

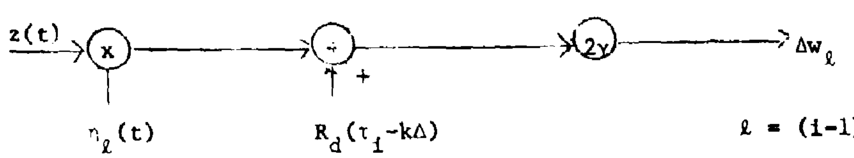
$$\underline{R}_{d\xi} = E\begin{bmatrix} d\xi_1 \\ d\xi_2 \\ \vdots \\ d\xi_{M+1} \\ \vdots \\ d\xi_{K(M+1)} \end{bmatrix} = E\begin{bmatrix} d(t) & s_1(t) \\ d(t) & s_2(t-\Delta) \\ \vdots & \\ d(t) & s_1(t-M\Delta) \\ \vdots & \\ d(t) & s_K(t-M\Delta) \end{bmatrix} = E\begin{bmatrix} d(t)d(t+\tau_1) \\ d(t)d(t+\tau_1-\Delta) \\ \vdots \\ d(t)d(t+\tau_1-M\Delta) \\ \vdots \\ d(t)d(t+\tau_K-M\Delta) \end{bmatrix} = \begin{bmatrix} R_d(\tau_1) \\ R_d(\tau_1-\Delta) \\ \vdots \\ R_d(\tau_1-M\Delta) \\ \vdots \\ R_d(\tau_K-M\Delta) \end{bmatrix} \tag{3.3-12}$$

In the above $R_d(\tau)$ is just the autocorrelation function of the desired signal d(t). For any given number of taps and their spacings, together with the known signal direction, $\underline{R}_{d\xi}$ can be completely specified if $R_d(\tau)$ is given.

Substituting (3.3-10) and (3.3-11) into (3.2-20), we obtain the desired algorithm to adjust the weight vector

$$\underline{W}_{j+1} = \underline{W}_j - 2\,\gamma_j\,z_j\,\underline{n}_j + 2\,\gamma_j\,\underline{R}_{d\xi} \tag{3.3-13}$$

During the training period, the information required to adjust the weights is just the signal autocorrelation function. z and $\underline{n}$ are available as real time functions. Algorithm (3.3-13) will be used extensively in designing an array processor. Its convergence properties are given in the next two sections. The implementation of this adaptive mechanism is very simple.



$\ell = (i-1)M+k$

A rather detailed structure is shown schematically in Fig. 7.

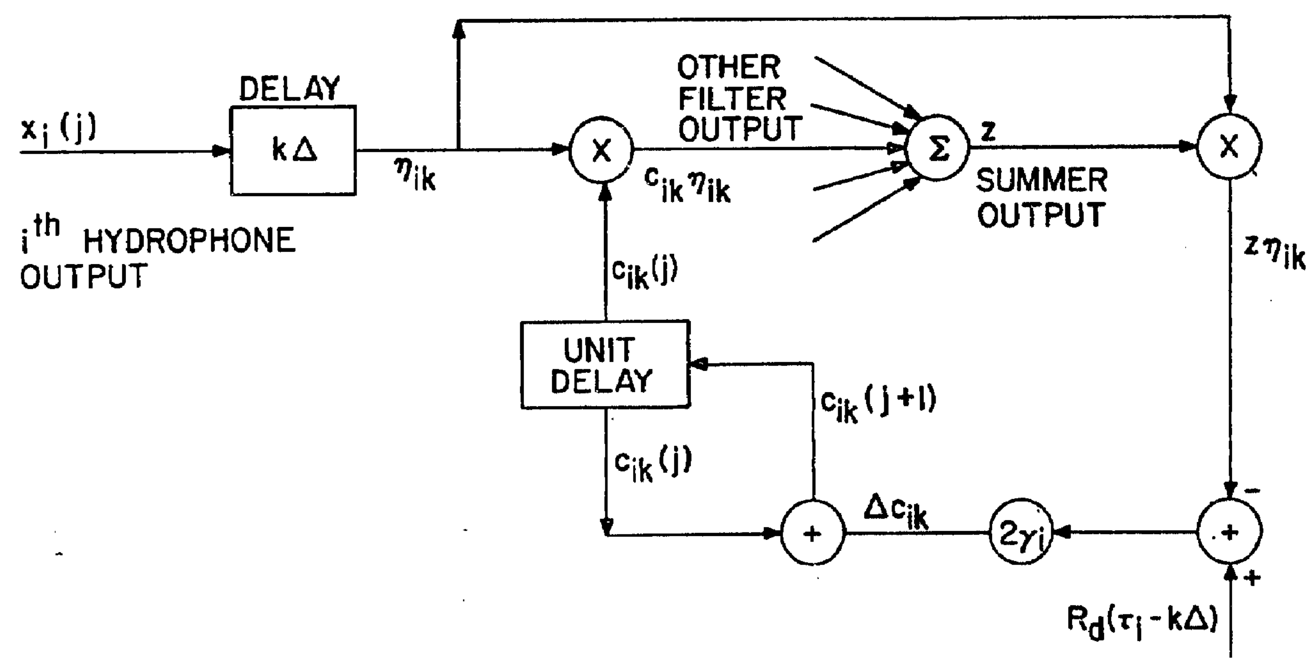


Figure 7. Adaptive Mechanism

$k = 0, 1, \ldots M$, Tap Index
$i = 1, 2, \ldots K$, Phone Index
$j = 1, 2, \ldots,$ Time Index

In comparing the formula for the optimum gains

$$\underline{W}_{op} = (\underline{R}_{\xi} + \underline{R}_{\nu})^{-1} \underline{R}_{d\xi}$$

with the recursive procedure

$$\underline{W}_{j+1} = \underline{W}_j + 2 \gamma_j \underline{R}_{d\xi} - 2 \gamma_j z_j \underline{n}_j$$

we see several advantages in using adaptive tapped-delay-line filters over non-adaptive tapped-delay-line filters:

(1) No noise field measurements are required since the weights are adjusted in the presence of normal hydrophone outputs.

(2) No solutions of simultaneous equations for the weights are required.

(3) When the signal correlation functions are used in (3.3-13) the difficulty of generating some simulated signals as proposed by Widrow and et al [16] is completely removed.

(4) It is not necessary to assume that the undesired interferences originate from point sources. The noise can take any realistic forms.

3.4 Physical Interpretations of the Convergence Conditions

It has been shown that algorithm (3.3-13)

$$\underline{W}_{j+1} = \underline{W}_j - 2 \gamma_j z_j \underline{n}_j + 2 \gamma_j \underline{R}_{d\xi} \qquad (3.4\text{-}1)$$

is derived from (3.2-20) with $\underline{c}$ replaced by $\underline{W}$, i.e.,

$$\underline{W}_{j+1} = \underline{W}_j - \gamma_j(\nabla Q_1 + \overline{\nabla Q_2}) \qquad (3.4\text{-}2)$$

Although (3.4-2) converges both in mean square and in probability under certain mathematical conditions, it is not clear whether these conditions can be met in reality. These conditions are repeated here for convenience.

(A) $\lim_{j\to\infty} \gamma_j = 0$, $\sum_{j=1}^{\infty} \gamma_j = \infty$, $\sum_{j=1}^{\infty} \gamma_j^2 < \infty$, $\gamma_j > 0$

(B)
$$\inf_{\varepsilon < ||\underline{W} - \underline{W}_{op}|| < \frac{1}{\varepsilon}} (\underline{W} - \underline{W}_{op})^T \nabla_w(\nabla Q_1 + \overline{\nabla Q_2}) > 0$$

(C) $E\{(\nabla Q_1 + \overline{\nabla Q_2})^T (\nabla Q_1 + \nabla Q_2)\}$

$$< d\,(\underline{W}_{op}^T \underline{W}_{op} + \underline{W}^T \underline{W}) \ , \ d > 0$$

The choice of γ_j which satisfies (A) is rather at our own disposal. We can always set $\gamma_j = \frac{\gamma}{j^\alpha}$ where $\gamma > 0$ and $1/2 < \alpha < 1$ to fulfill the requirements of (A). The remaining conditions depend on the surface of the error gradient, which in turn depends on the choice of error criterion and the physical system under consideration.

We shall show in the following two lemmas that conditions (B) and (C) are satisfied if (1) the error function $Q(e)$ is strictly convex; (2) the second derivative of $Q(e)$ with respect to e exists and is uniformly bounded; (3) all signals (useful signal, ambient noise and interferences) are generated from physically realizable sources and thus their second order statistics are uniformly bounded. The first two conditions are definitely satisfied because the performance criterion employed here is just the mean squared error so that $Q(e) = e^2$, which is strictly convex and $\partial^2 Q/\partial e^2 = 2$ is uniformly bounded. The third condition concerning the boundedness of the correlation functions of the input processes is also satisfied in most practical situations.

Consequently, we can conclude that all the convergence conditions can be met in practice and the adaptive schemes should be operative in adjusting the weights on the tapped-delay lines.

Lemma 1 : For the tapped delay line filters, if $Q(e) = e^2$ then at the neighborhood of $\underline{W}_{op}$ minimizing $E\{Q(e)\}$ the following statement is true:

$$\inf_{\varepsilon < ||\underline{W} - \underline{W}_{op}|| < \frac{1}{\varepsilon}} (\underline{W} - \underline{W}_{op})^T E\{\nabla Q_1 + \overline{\nabla Q_2}\} > 0$$

$\varepsilon > 0$

Proof: Since $I = E[Q]$ has a minimum at $\underline{W} = \underline{W}_{op}$, we can write for $k = 1, 2, \ldots K(M+1)$

$$\frac{\partial I}{\partial w_k} \gtreqless 0 \quad \text{for} \quad w_k \gtreqless w_{op}^{(k)} \tag{3.4-3}$$

thus

$$(w_k - w_{op}^{(k)}) \frac{\partial I}{\partial w_k} \geq 0 \quad \text{for all} \quad k \tag{3.4-4}$$

and

$$\inf_{\varepsilon < ||\underline{W} - \underline{W}_{op}|| < \frac{1}{\varepsilon}} (\underline{W} - \underline{W}_{op})^T E\{\nabla Q_1 + \overline{\nabla Q_2}\} > 0 \tag{3.4-5}$$

Lemma 2 :

Let $Q = Q_1 + Q_2 = e^2$. If the second order statistics of the input processes are hounded, then for the tapped-delay-line filters under study the following condition is always satisfied:

$$E\{(\nabla Q_1 + \overline{\nabla Q_2})^T(\nabla Q_1 + \overline{\nabla Q_2})\} \leq k_1 (\underline{W}_{op}^T \underline{W}_{op} + \underline{W}^T \underline{W}), \; k_1 > 0$$

Proof

Using a Taylor series expansion about $\underline{W} = \underline{W}_{op}$, we have

$$\nabla Q(\underline{W}) = \nabla Q(\underline{W}) \Big|_{\underline{W} = \underline{W}_{op}} + \underline{J} \Big|_{\underline{W} = \underline{W}_{op}} (\underline{W} - \underline{W}_{op}) \tag{3.4-6}$$

where $\underline{J}$ is the Jacobian having elements

$$J_{ik} = \frac{\partial^2 Q}{\partial w_i \partial w_k} \; ; \; i, \; k = 1,2,\ldots, K(M+1) \tag{3.4-7}$$

Since the error function is

$$e(t) = d(t) - z(t) = d(t) - \sum_{k=1}^{K(M+1)} w_k \eta_k(t) \tag{3.4-8}$$

we see that

$$\frac{\partial Q}{\partial w_k} = \frac{\partial Q}{\partial e} \frac{\partial e}{\partial w_k} = \frac{\partial Q}{\partial e} [- \eta_k(t)] \tag{3.4-9}$$

and $$\frac{\partial^2 Q}{\partial w_k \partial w_\ell} = \frac{\partial^2 Q}{\partial e^2} \eta_k(t) = 2\, \eta_k(t)\, \eta_\ell(t) \tag{3.4-10}$$

Therefore, the averaged value of the error gradient can be written in the following form in view of the above expressions

$$\overline{\nabla Q} = \overline{\nabla Q}\Big|_{\underline{W} = \underline{W}_{op}} + 2\, \underline{R}_\eta (\underline{W} - \underline{W}_{op}) = 2\, \underline{R}_\eta (\underline{W} - \underline{W}_{op}) \tag{3.4-11}$$

because $\overline{\nabla Q} = 0$ at the optimum point and $\underline{R}_\eta$ is the input correlation matrix with elements $\overline{\eta_k(t)\eta_\ell(t)}$ for $k, \ell = 1,2,\ldots, K(M+1)$.

Note that

$$E\{(\nabla Q_1 + \overline{\nabla Q_2})^T(\nabla Q_1 + \overline{\nabla Q_2})\}$$

$$\leq E\{\nabla^T Q_1 + \overline{\nabla^T Q_2}\}\, E\{\nabla Q_1 + \overline{\nabla Q_2}\} \tag{3.4-12}$$

and for real variables

$$a^2 + b^2 > -2ab \quad \text{from} \quad (a+b)^2 > 0$$

$$(a-b)^2 = a^2 + b^2 \leq 2(a^2 + b^2) \tag{3.4-13}$$

The desired result is obtained by substituting Eq. (3.4-12) into Eq.(3.4-11) and setting a constant

$$k_1 = \sup_{\text{all } k,\ell} \left|\overline{\eta_k(t)\eta_\ell(t)}\right| \tag{3.4-14}$$

for all k and ℓ. The constant k_1 defined above will be bounded if the second order statistics of the input processes are bounded.

3.5 Convergence Properties of the Adaptive Tapped-Delay-Line Filters

Having found an algorithm which converges in some sense, we shall now investigate how fast it converges. In other words, we would like to know how fast the parameters approach to their values and the mean-squared-error at each stage during the adaptation period. The effect of the input statistics on the

rate of convergence will be determined. The adaptive behavior adjusted in the presence or absence of the target signal will also be studied.

Rewrite algorithm (3.3-13) here for convenience

$$\underline{W}_{j+1} = \underline{W}_j + 2\,\gamma_j\,\underline{R}_{d\xi} - 2\,\gamma_j\,z_j\,\underline{n}_j \tag{3.5-1}$$

Since the summer output is

$$z_j = \underline{W}_j^T\,\underline{n}_j \tag{3.5-2}$$

we have

$$\underline{W}_{j+1} = (1 - 2\,\gamma_j\,\underline{n}_j\,\underline{n}_j^T)\,\underline{W}_j + 2\,\gamma_j\,\underline{R}_{d\xi} \tag{3.5-3}$$

Taking the mathematical expectation of (3.5-3) and diagonizing the input correlation matrix $\underline{R}_n$ such that

$$\underline{R}_n = \underline{P}^{-1}\,\underline{\Lambda}\,\underline{P} \tag{3.5-4}$$

we obtain

$$E[\underline{W}_{j+1}] = (1 - 2\gamma_j\,\underline{P}^{-1}\,\underline{\Lambda}\,\underline{P})\,E[\underline{W}_j] + 2\gamma_j\,\underline{R}_{d\xi} \tag{3.5-5}$$

where $\underline{P}$ is an orthonormal matrix and $\underline{\Lambda}$ is the corresponding eigenvalue matrix.

Some comments are in order

1) In Eq. (3.5-4) the input correlation matrix assumes different values depending upon whether the input contains noise only or signal plus noise. When both the target signal and the undesired noise are present, the output of the i^{th} hydrophone is

$$x_i(t) = s_i(t) + n_i(t) \tag{3.5-6}$$

so that the various delayed inputs $n_k(t)$ contain signal components $\xi_k(t)$ as well as noise components $\nu_k(t)$

$$n_k(t) = \xi_k(t) + \nu_k(t)\ ,\ k = 1, 2, \ldots, K(M+1) \tag{3.5-7}$$

and

$$\underline{n}\, z = (\underline{\xi} + \underline{\nu})\ (\underline{\xi} + \underline{\nu})^T\ \underline{W} \qquad (3.5\text{-}8)$$

$$E[\underline{n}\ z] = E[\underline{\xi}\ \underline{\xi}^T + \underline{\nu}\ \underline{\nu}^T]\ E[\underline{W}] = (\underline{R}_\xi + \underline{R}_\nu)\ E[\underline{W}] \qquad (3.5\text{-}9)$$

$\underline{R}_\xi$ and $\underline{R}_\nu$ are the input signal correlation and input noise correlation matrices. Thus, it is important to keep in mind that

$$\underline{R}_\eta = \underline{R}_\xi + \underline{R}_\nu \quad \text{when} \quad x_i(t) = s_i(t) + n_i(t) \qquad (3.5\text{-}10)$$

and

$$\underline{R}_\eta = \underline{R}_\nu \quad \text{when} \quad x_i(t) = n_i(t) \qquad (3.5\text{-}11)$$

2) In taking the average over Eq. (3.5-8) it is assumed that $\underline{W}$ is statistically independent of $\underline{n}$. Although $\underline{W}$ cannot affect $\underline{n}$ in any manner, the increment of $\underline{W}$ at each stage is related to $\underline{n}$ by Eq.(3.5-1). Since the increment is generally very small and the total effect involves addition of a large number of small increments, we can assume $E[\underline{n}\ \underline{n}^T\underline{W}] = E[\underline{n}\ \underline{n}^T]\ E[\underline{W}]$ in a manner similar to that used in the analysis of phase-locked loops[1].
Thus for large j (at later stages during the training period) there should be little correlation between $\underline{W}_{j+1}$ and $\underline{n}_{j+1}$.

In returning to Eq. (3.5-5), let us define a new weight vector

$$\underline{W}' = \underline{P}\ \underline{W} \qquad (3.5\text{-}12)$$

and a new delayed input vector

$$\underline{n}' = \underline{P}\ \underline{n} \qquad (3.5\text{-}13)$$

Since[2] $\underline{R}_{d\xi} = \underline{R}_\eta\ \underline{W}_{op}$ as seen from Eq. (2.3-25),
we transform Eq. (3.5-5) into

$$E[\underline{W}'_{j+1}] = (1 - 2\gamma_j\ \underline{\Delta})\ E[\underline{W}'_j] + 2\gamma_j\ \Delta\ \underline{W}'_{op} \qquad (3.5\text{-}14)$$

---

1 See A. J. Viterbi, Principles of Coherent Communication McGraw Hill Book Co., N. Y., 1966.

2 The optimum weight vector $\underline{W}_{op}$ assumes $\underline{R}_\eta^{-1}\ \underline{R}_{d\xi}$ or $\underline{R}_\nu^{-1}\ \underline{R}_{d\xi}$ depending on the training environment.

or

$$E[\underline{W}_{j+1}] - \underline{W}'_{op} = (1 - 2\gamma_j \underline{\Lambda}) \; (E[\underline{W}'_j] - \underline{W}'_{op}) \tag{3.5-15}$$

Now consider any particular component of $\underline{W}'$, and for clarity no subscript or superscript denoting the component is used. Then we obtain a difference equation for $E[w'_j] = \bar{w}'_j$

$$\bar{w}'_{j+1} - w'_{op} = (1 - 2\gamma_j \; \lambda) \; (\bar{w}'_j - w'_{op}) \tag{3.5-16}$$

whose solution is

$$\bar{w}'_{j+1} = (w'_1 - w'_{op}) \prod_{k=1}^{j} (1 - 2\gamma_j \; \lambda) + w'_{op} \tag{3.5-17}$$

We shall now calculate the mean square of the weights.

If we first take the outer product and then the mathematical expectation on both sides of Eq. (3.5-1), we can write after some algebraic manipulations[1] [17]

$$\begin{aligned}\overline{\underline{W}_{j+1} \underline{W}^T_{j+1}} = \overline{\underline{W}_j \underline{W}^T_j} + 4\gamma_j \{\underline{R}_n \overline{(\underline{W}_{op} - \underline{W})\underline{W}^T}\}^s_j \\ + 4\gamma_j^2 \overline{e^2_j \underline{n}_j \underline{n}^T_j}\end{aligned} \tag{3.5-18}$$

where $\{A\}^s$ denotes the symmetric part of matrix A and

$$\{A \; B\}^s = \frac{1}{2} (\underline{A} \; \underline{B}^T + \underline{B} \; A^T)$$

For large j , the following approximation can be made

$$\overline{e^2_j \underline{n}_j \underline{n}_j} \simeq \overline{e^2_{min}} \; \overline{\underline{n}_j \underline{n}_j} = \overline{e^2_{min}} \; \underline{R}_n \tag{3.5-19}$$

which can be viewed as a Taylor series expansion around the optimum point and with higher order terms neglected.

---

[1] This is done by combining the following steps:

a. $\underline{W}_{j+1} \underline{W}^T_{j+1} = \underline{W}_{j+1} \underline{W}_{j+1} + 2\gamma_j \; e_j(\underline{W}_j \underline{n}^T_j + \underline{n}_j \underline{W}^T_j) + 4\gamma_j^2 e^2_j \underline{n}_j \underline{n}^T_j$

b. $\overline{e_j(\underline{W}_j \underline{n}^T_j + \underline{n}_j \underline{W}^T_j)} = 2 \{\underline{R}_n(\underline{W}_{op} - \underline{W}_j) \underline{W}^T_j\}^s$

c. $\{ \; \}^s$ is used to make the expression compact and the superscript can be removed in dealing with diagonal terms of a square matrix.

Therefore, Eq. (3.5-18) becomes

$$\underline{W}_{j+1}\,\underline{W}_{j+1}^T \approx \underline{W}_j\,\underline{W}_j^T + 4\,\gamma_j\,\{\underline{R}_\eta\,(\underline{W}_{op} - \underline{W})\underline{W}^T\}_j^s$$

$$+ 4\,\gamma_j^2\,\overline{e_{min}^2}\,\underline{R}_\eta \tag{3.5-20}$$

Using the transformation $\underline{W}' = \underline{P}\,\underline{W}$ defined in Eq. (3.5-12), we can express the diagonal terms of the above matrix as

$$\overline{\{\underline{W}_{j+1}'\,\underline{W}_{j+1}'^T\}^D} \approx \overline{\{\underline{W}_j'\,\underline{W}_j'^T\}^D} + 4\,\gamma_j\,\overline{\{\underline{\ }\,(\underline{W}_{op}' - \underline{W}')\underline{W}'^T\}_j^D}$$

$$+ 4\,\gamma_j^2\,e_{min}^2\,\underline{\ } \tag{3.5-21}$$

while the outer product of Eq. (3.5-14) is

$$\overline{\underline{W}_{j+1}'}\,\overline{\underline{W}_{j+1}'}^T = \{(1 - 4\,\gamma_j\,\underline{\ })\overline{\underline{W}_j'}\,\overline{\underline{W}_j'}^T\}^s + 4\,\gamma_j\,\underline{W}_{op}'\,\overline{\underline{W}_j'}^T \tag{3.5-22}$$

Let

$$\underline{V}_j = \overline{(\underline{W}_j' - \overline{\underline{W}_j'})(\underline{W}_j' - \overline{\underline{W}_j'})^T} = \overline{(\underline{W}'\,\underline{W}'^T)}_j - \overline{\underline{W}_j'}\,\overline{\underline{W}_j'}^T \tag{3.5-23}$$

Subtracting the diagonal terms of Eq. (3.5-21) from those of Eq. (3.5-22) yields

$$\underline{V}_{j+1}^D = (1 - 4\,\gamma_j\,\underline{\Lambda})\,\underline{V}_j^D + 4\,\gamma_j^2\,\underline{\Delta}\,e_{min}^2 \tag{3.5-24}$$

which has elements of the form $\overline{(w_j')^2} - (\overline{w_j'})^2$.

Thus, for any particular component of $\underline{V}_j^D$, we have

$$\overline{(w_{j+1}')^2} - (\overline{w_j'})^2 = v_{j+1} = (1 - 4\,\gamma_j\,\lambda)\,v_j + 4\,\gamma_j^2\,\lambda\,\overline{e_{min}^2} \tag{3.5-25}$$

Iterating backward,

$$v_{j+1} = v_1 \prod_{k=1}^{j} (1 - 4\,\gamma_k\,\lambda) + 4\,\lambda\,\overline{e_{min}^2} \sum_{k=1}^{j} \gamma_k^2 \prod_{\ell=k+1}^{j} (1 - 4\,\gamma_\ell\lambda) \tag{3.5-26}$$

But $v_1 = 0$ because $(\overline{w_1'}) = w_1'$, so that

$$v_{j+1} = \overline{(w_{j+1}' - \overline{w}_{j+1}')^2} = 4\ \lambda\ \overline{e^2_{min}} \sum_{k=1}^{j} \gamma_k^2 \prod_{\ell=k+1}^{j} (1 - 4\ \gamma_\ell\ \lambda) \tag{3.5-27}$$

Using Eqs. (3.5-17) and (3.5-27) we shall find the rate of convergence in terms of weight variance and the mean squared error.

The weight variance is

$$\overline{||\underline{W}_{j+1} - \underline{W}_{op}||^2} = \sum_{m=1}^{K(M+1)} \overline{[w_{j+1}^{(m)} - w_{op}^{(m)}]^2} \tag{3.5-28}$$

and the mean squared error at each stage of adaptation is from Eq. (2.3-28)

$$\overline{e^2_{j+1}} - \overline{e^2_{min}} = (\underline{W}_{j+1} - \underline{W}_{op})^T\ \underline{R}_n\ (\underline{W}_{j+1} - \underline{W}_{op})$$

$$= (\underline{W}_{j+1}' - \underline{W}_{op}')\ \Delta\ (\underline{W}_{j+1}' - \underline{W}_{op}') \tag{3.5-29}$$

where Δ and $\underline{W}'$ are defined by Eq. (3.5-4) and Eq. (3.5-12) respectively. The expected difference between the mean squared error at each stage during the training period and the minimum mean squared error is then

$$E[\overline{e^2_{j+1}} - \overline{e^2_{min}}] = E\left\{ \sum_{m=1}^{K(M+1)} \lambda_m\ (w_{j+1}'^{(m)} - w_{op}'^{(m)})^2 \right\}$$

$$= \sum_{m=1}^{K(M+1)} \lambda_m\ \overline{(w_{j+1}'^{(m)} - w_{op}'^{(m)})^2} \tag{3.5-30}$$

Since

$$\overline{(w_{j+1}' - w_{op}')^2} = \overline{(w_{j+1}' - \overline{w}_{j+1}')^2} + (\overline{w}_{j+1}' - w_{op}')^2 \tag{3.5-31}$$

the weight variance is

$$\overline{||\underline{W}_{j+1} - \underline{W}_{op}||^2} = \sum_{m=1}^{K(M+1)} \left\{ 4\ \lambda_m\ \overline{e^2_{min}} \sum_{k=1}^{j} \gamma_k^2 \prod_{\ell=k+1}^{j} (1 - 4\gamma_\ell\lambda_m) \right.$$

$$\left. + [(w_1'^{(m)} - w_{op}'^{(m)}) \prod_{k=1}^{j} (1 - 2\ \gamma_k\ \lambda_m)]^2 \right\} \tag{3.5-32}$$

and the mean squared error becomes

$$\overline{e^2_{j+1}} - \overline{e^2_{min}} = \sum_{m=1}^{K(M+1)} 4\ \lambda_m^2\ \overline{e^2_{min}} \sum_{k=1}^{j} \gamma_k^2 \prod_{\ell=k+1}^{j} (1 - 4\ \gamma_\ell \lambda_m)$$

$$+ \sum_{m=1}^{K(M+1)} \lambda_m\ (w_1'^{(m)} - w_{op}'^{(m)})^2 \left(\prod_{k=1}^{j}(1 - 2\ \gamma_k \lambda_m)\right)^2 \qquad (3.5\text{-}33)$$

Eqs. (3.5-32) and (3.5-33) are the desired expressions for the rate of convergence. Although they appear to be quite complicated, simpler results can be achieved by setting the weighting sequence γ_j to some special forms.

a) $\gamma_j = \dfrac{1}{2(j+1)\lambda}$ (3.5-34)[1]

This really means that γ_j is a diagonal matrix whose m^{th} element is $\dfrac{1}{2(j+1)\lambda_m}$ for $m = 1,2,\ldots K(M+1)$. Eq. (3.5-34) is a legitimate choice as it satisfies all the required conditions for convergence.

Since

$$\prod_{k=1}^{j}(1 - 2\ \gamma_k\ \lambda) = \prod_{k=1}^{j}\ (1 - \frac{1}{k+1}) = \frac{1}{j+1} \qquad (3.5\text{-}35)$$

and [see (C-8) of Appendix C]

$$\prod_{\ell=k+1}^{j}(1 - 4\ \gamma_\ell \lambda) = \prod_{\ell=k+1}^{j}(1 - \frac{2}{\ell+1}) \simeq \frac{(k+1)^2}{(j+1)^2} \qquad (3.5\text{-}36)$$

the mean squared error at each stage is

$$\overline{e^2_{j+1}} - \overline{e^2_{min}} \simeq \frac{j}{(j+1)^2} \sum_{m=1}^{K(M+1)} \overline{e^2_{min}} + \frac{1}{(j+1)^2} \sum_{m=1}^{K(M+1)} \lambda_m\ (w_1'^{(m)} - w_{op}'^{(m)})^2$$

$$= K(M+1)\ \overline{e^2_{min}}\ \frac{j}{(j+i)^2} + \frac{1}{(j+1)^2}\ (\underline{W}_1 - \underline{W}_{op})^T\ \underline{R}_\eta (\underline{W}_1 - \underline{W}_{op})$$

(3.5-37)

---

[1] This is the optimum choice of γ_j which provides the fastest rate of convergence and can be derived by making $\overline{e^2_j}$ a minimum for each $j=1,2,\ldots$. Simulation results also confirm this argument.

and the weight variance is

$$\overline{\|\underline{W}_{j+1} - \underline{W}_{op}\|^2} \approx \frac{j}{(j+1)^2} \overline{e^2_{min}} \sum_{m=1}^{K(M+1)} \frac{1}{\lambda_m}$$

$$+ \frac{1}{(j+1)^2} \overline{\|\underline{W}_1 - \underline{W}_{op}\|^2} \tag{3.5-38}$$

The significance of the above expressions is apparent.

The initial error at time $j = 1$ is just the norm of the difference between the initial and optimum gains in the parameter space, i.e.,

$$\overline{e^2_1} = \overline{e^2_{min}} + (\underline{W}_1 - \underline{W}_{op})^T \underline{R}_\eta (\underline{W}_1 - \underline{W}_{op}) \tag{3.5-39}$$

The initial mean-squared error will decrease at the rate of $1/j^2$, j being the adaptation time. However, the first term on the right-hand side of Eq. (3.5-37) decreases at the rate of $j/(j+1)^2 \approx 1/j$ and will definitely dominate the first term during the later training period. It is proportional to the total number of gains being adjusted and the absolutely obtainable mean-squared error using that many taps. That is, for large j

$$\overline{e^2_{j+1}} - \overline{e^2_{min}} \approx \frac{1}{j} K(M+1) \overline{e^2_{min}} \tag{3.5-40}$$

It appears that more training time would be required to make $\overline{e^2_j}$ approach to $\overline{e^2_{min}}$ if more weights [K(M+1) in number] are to be adjusted. Actually, $\overline{e^2_{min}}$ is a monotone decreasing function of M . It was shown in Chapter II that

$$\overline{e^2_{min}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 \, d\omega - \sum_{m=1}^{K(M+1)} w_m^2 \tag{3.5-41}$$

where $H(\omega)$ is the continuous transfer function to be approximated. Since this quantity does not decrease linearly with M , there is always compromise to be made between the training time and the accuracy of approximation. Therefore we should keep in mind that using too many taps may do more harm (longer training

period) than good (smaller mean-squared error). That the mean squared error decreases at the rate of $\frac{1}{j^\alpha}$ for $\gamma_j = (\frac{1}{j^\alpha})$ is a well-known fact in employing the methods of stochastic approximation. Here, however, we have shown explicitly the dependence of error rate on system parameters such as the number of hydrophones in an array, number of taps for each individual filter, and the input statistics. As a simple illustration, suppose that we want to reduce the error in a single filter to about 1% of its minimum

$$(\overline{e^2_{j+1}} - \overline{e^2_{min}})/\overline{e^2_{min}} = 0.01$$

then for 100 taps we need roughly J = 10,000 samples to adjust the weights or equivalently about 10 seconds of real-time data for a sampling rate of 1,000 samples per second.

The time required to make this same adjustment in an array of filters need not be much greater since the adjustment can be done in parallel processors operating simultaneously. Parallel processing is quite feasible here, since the basic algorithm is so simple.

b) $\gamma_j = \frac{1}{2(j+1)}$ (3.5-42)

The choice of γ_j defined in Eq. (3.5-34) requires some a priori knowledge about the input statistics. If the noise correlation matrix is unknown, an assumption forcing us to apply adaptive techniques, the input correlation matrix and thus the eigenvalues cannot be determined. Here we shall consider an arbitrary sequence $\gamma_j = 1/2(j+1)$.

Since (see Appendix C)

$$\prod_{k=1}^{j} (1 - \frac{\lambda}{j+1}) = \frac{\Gamma(j+2 - \lambda)}{(j+1)!\ \Gamma(2-\lambda)} \approx \frac{1}{\Gamma(2-\lambda)(j+1)^\lambda} \quad (3.5\text{-}43)$$

for $j >> 1$ and $j >> \alpha$

and

$$\prod_{\ell=k+1}^{j} \left(1 - \frac{2\lambda}{j+1}\right) \simeq \frac{(k+1)^{2\lambda}}{(j+1)^{2\lambda}} \tag{3.5-44}$$

Eq. (3.5-33) becomes

$$\overline{e_{j+1}^2} - \overline{e_{min}^2} \simeq \sum_{m=1}^{K(M+1)} \left\{ \frac{\lambda_m^2 \overline{e_{min}^2}}{(j+1)^{2\lambda_m}} \sum_{k=1}^{j} (k+1)^{2\lambda_m} + \frac{\lambda_m}{(j+1)^{2\lambda_m}} \left(\frac{w_1'^{(m)} - w_{op}'^{(m)}}{\Gamma(2-\lambda_m)} \right)^2 \right\} \tag{3.5-45}$$

which of course reduces to Eq. (3.5-37) when $\lambda_m = 1$ for $m = 1,2,\ldots, K(M+1)$.

c) $\gamma_j = \gamma = \text{constant}$ (3.5-46)

The expressions for γ_j defined by Eqs. (3.5-34) and (3.5-42) satisfy the conditions for convergence as stated in Section 3.2. In these cases the γ's and thus the gain increment $\Delta \underline{W}_j = \underline{W}_{j+1} - \underline{W}_j$ become smaller and smaller as time j proceeds during the adaptation period. It is anticipated that the rate of convergence will be increased if a small constant value is set for γ. As shown by Comer [42], the algorithm with constant γ has comparatively little noise resistance. Furthermore, in the presence of measuring error with variance σ^2, convergence in the usual sense does not occur, but

$$\lim_{j\to\infty} \overline{||\underline{W}_{j+1} - \underline{W}_{op}||^2} < F(\gamma , \sigma^2) \tag{3.5-46}$$

and

$$F(\gamma , \sigma^2) \to 0 \quad \text{as} \quad \gamma \to 0 \tag{3.5-47}$$

We shall next study the rate of convergence when γ is a constant. From Eq. (3.5-17) we see that with $\gamma_j = \gamma = \text{constant}$

$$\begin{aligned} \overline{w_{j+1}'} &= (w_1' - w_{op}') \prod_{k=1}^{j} (1 - 2\gamma\lambda) + w_{op}' \\ &= (1 - 2\gamma\lambda)^j (w_1' - w_{op}') + w_{op}' \end{aligned} \tag{3.5-48}$$

Since

$$a + a\gamma + a\gamma^2 + \cdots + a\gamma^{n-1} = \frac{a(1 - \gamma^n)}{1 - \gamma} \tag{3.5-49}$$

We can obtain

$$\sum_{k=1}^{j} (1 - 4\gamma\lambda)^{-k} = \frac{1}{4\gamma\lambda} [(1 - 4\gamma\lambda)^{-(j-1)} - 1] \tag{3.5-50}$$

Thus,

$$\sum_{k=1}^{j} \gamma_k^2 \prod_{\ell=k+1}^{j} (1 - 4\gamma\lambda) = \sum_{k=1}^{j} \gamma^2 (1 - 4\gamma\lambda)^{j-k-1}$$

$$= \gamma^2 (1 - 4\gamma\lambda)^{j-1} \sum_{k=1}^{j} (1 - 4\gamma\lambda)^{-k}$$

$$= \gamma^2 \frac{1}{4\gamma\lambda} [1 - (1 - 4\gamma\lambda)^{j-1}] \tag{3.5-51}$$

and Eq. (3.5-27) becomes

$$\overline{(w'_{j+1} - \overline{w'_{j+1}})^2} = \overline{e^2_{min}}\, \gamma\, [1 - (1 - 4\gamma\lambda)^{j-1}] \tag{3.5-52}$$

The mean-squared error is then

$$\overline{e^2_{j+1}} - \overline{e^2_{min}} = \overline{e^2_{min}}\, \gamma \sum_{m=1}^{K(M+1)} \lambda_m [1 - (1 - 4\gamma\lambda_m)^{j-1}]$$

$$+ \sum_{m=1}^{K(M+1)} \lambda_m (w_1'^{(m)} - w_{op}'^{(m)})^2 (1 - 2\gamma\lambda_m)^{2j} \tag{3.5-53}$$

It is seen from the above expression that if the error is to decrease at all, one basic requirement should be met, i.e.,

$$0 < 1 - 4\gamma\lambda_m < 1 \quad \text{with} \quad \gamma > 0 \tag{3.5-54}$$

$$\text{for} \quad m = 1, 2, \ldots, K(M+1)$$

which implies

$$0 < \gamma < \frac{1}{4\lambda_{max}} \tag{3.5-55}$$

λ_{max} is the largest eigenvalue of the correlation matrix $\underline{R}_{\eta}$. Thus γ = constant cannot be set at will if stability of the adaptive loop is to be maintained.

Conclusion

The adjustable gains under the operation of our proposed adaptive scheme using signal correlation functions converge to two different sets of optimum values depending on whether the input contains target signal or not during the training period, i.e.,

$$\lim_{j\to\infty} \underline{W}_j = (\underline{R}_{\xi} + \underline{R}_{\nu})^{-1} \underline{R}_{d\xi} \quad \text{if} \quad \underline{x} = \underline{s} + \underline{n}$$

$$\lim_{j\to\infty} \underline{W}_j = \underline{R}_{\nu}^{-1} \underline{R}_{d\xi} \quad \text{if} \quad \underline{x} = \underline{n} \ .$$

The mean squared error decreases approximately as the first power of the adaptation time. The rate of convergence is essentially indifferent to the number of weights to be adjusted as our algorithm allows simultaneous adjustments. The size of error, on the other hand, does depend on the total number of taps and the difference between the initial and the optimum values of the weights. It is also of importance to note that the weighting sequence cannot be selected at will. Although $\gamma_j = \frac{\gamma}{j^{\alpha}}$, $\frac{1}{2} < \alpha \leq 1$ satisfies all the conditions for convergence for any positive constant γ , this constant should not exceed $\frac{1}{4\lambda_{max}}$ if stability of the adaptive loop is to be maintained. This is especially important during the early stages of adaptation. Simulation results are given in Chapter six.

3.6 Further Remarks on the Operations of the Proposed System

a) Choice of the Initial Weights

Although the adjustable weights can be set to any values at the beginning of the adaptive process, it is desirable to set them not too far from their optimum

by using whatever information is available concerning the statistics of the noise field. The formula for calculating the optimum gains can be utilized to start the initial computation with inaccurate noise statistics. This kind of choice will shorten the adaptation period and thus reduce the cost of operation. In cases where absolutely no such information is known, the gains associated with the input delayed by τ_i ($i = 1,2,\ldots K$) are set to 1 and the rest to zero so that a square law detector is used at the starting moments. As the adaptive proceeds, the whole system will gradually be transformed into an optimum one. Any target signal not detectable during the early stages can probably be ferretted out at a later time.

b) Problem of Signal Suppression

In most adaptive detection systems such as those of Glaser [18], Jackowatz, etc. [19] more errors are made as the input signal-to-noise ratio is decreased. In fact, it has been hypothesized that if the signal-to-noise ratio is gradually decreased, eventually a point is reached where instability occurs, with consequent breakdown of the system. That is, for signal-to-noise ratios below a certain level, the number of errors degrades the quality of the measurements to the extent that the use of the erroneous measurements by the detection results in even more errors. This, in turn, causes even poorer measurements, and so on until a complete collapse of the system performance to an error rate of one-half occurs. In our system, however, adaptation always takes place regardless of the presence of the target signal. It is therefore reasonable to anticipate that there will be no signal suppression phenomenon.

c) Problem of Uncertain Signal Power

In designing most non-adaptive optimum detection systems, complete statistical knowledge is required for both the signal and the noises. That is, their spectral shapes as well as their power levels are assumed to be known. The

proposed adaptive system assumes no information about the noise fields, which represents one of major advantages in applying iterative procedures. Although it is reasonable to assume that the general shape of signal spectrum is known, signal power level may be in some cases uncertain before detection. To get more insight about the operation of the proposed system one may ask how the uncertain signal power affects the system performance. This can be answered by studying what algorithm (3.5-1) converges to if signal is indeed present in the postulated direction but has a power level different from that assumed.

It is shown in Appendix D that if the assumed signal power differs from the actual power by a multiplicative constant, the gains adjusted according to algorithm (3.5-1) will converge in mean as well as in mean square to their optimum values multiplied by the same constant. Consequently, the asymptotic structure of the proposed system will differ from the optimum one by a multiplicative constant if incorrect signal power is assumed during the adaptation period. For a fixed threshold the detectability of the detector, in terms of false alarm rate and miss probability, will be degraded to an extent depending on how the constant deviates from unity. The threshold should be adjusted around its normal operating level as a function of the signal power. If, however, some kind of display device is available as in most practical cases to observe the directivity pattern of the array system, uncertain signal power will not affect the sensitivity of the pattern. The output signal-to-noise ratio remains essentially unchanged.

CHAPTER FOUR

ADAPTATION IN A NONSTATIONARY ENVIRONMENT

4.1 Introduction

In Chapter III iterative procedures were derived which involve no noise statistics and no explicit time-averaging. It is generally noted that the optimum processor can do significantly better than the conventional processor for highly directional noise. However, highly directional noise fields are likely to be nonstationary. For example, in the sonar array problem, the most likely sources of directional noise or interference are ships, and ships may be moving. Under this situation the input covariance matrix and hence the optimum gains on the tapped-delay lines will be a function of time. It is obviously desirable to modify the algorithms in such a way that adaptation can still be accomplished in a nonstationary environment. Otherwise, if we still use the same algorithm to estimate the gains, the actual optimum point in the parameter space would have moved to some other place before a steady state is reached. This is a very important problem frequently encountered in practice.

In this chapter we shall consider several partial solutions to this difficult problem. These solutions are partial because each one of them can be applied to very restrictive cases under particular assumptions. If the law governing the parameter variation is known completely, we can generalize the dynamic stochastic approximation method [36] to adjust the time-varying parameters. In case the dynamics of parameter variations is generated by a special mechanism and some pertinent statistics are available, we can then apply the Kalman filtering techniques to this nonstationary problem. Cases mostly encountered in practice are nevertheless different from these two. We cannot expect to know the equation of parameter variation exactly, nor do we have complete statistics. If all the information we have about the noise fields is the rate of change, we shall just use the ordinary procedure and determine the effect of nonstationarity on its

convergence properties.

4.2 Application of the Method of Dynamic Stochastic Approximation

When the random environment is nonstationary with time-varying statistics, the optimum parameter set $\underline{\theta} = \underline{W}_{op}$ becomes a function of time index j. Its value at time j will be denoted by $\underline{\theta}_j$. It is assumed here that the law governing the variation is known, although the sequence to be estimated is unknown. In this case the generalized dynamic stochastic approximation method developed in Appendix E can be applied to the design of adaptive tapped-delay line filters. If the variation of $\underline{\theta}$ is governed by a known operator L such that $\underline{\theta}_{j+1} = L(\underline{\theta}_j, j)$ then the desired algorithm is given by (Eq. E-19)

$$\underline{W}_{j+1} = L(\underline{W}_j, j) - \gamma_j \nabla Q(\underline{x}_j | \underline{W}_j) \tag{4.2-1}$$

where the C's are replaced by the W's. Upon using (4.2-1) the adjustment procedure for our delay line filters becomes

$$\underline{W}_{j+1} = L(\underline{W}_j, j) + 2\gamma_j \underline{R}_{d\xi} - 2\gamma_j z_j \underline{n}_j \tag{4.2-2}$$

The above formulation is restricted to the case where the dynamics of the optimum set $\underline{\theta}$ are described by a homogeneous difference equation

$$\underline{\theta}_{j+1} = \underline{L}_{j+1,j} \underline{\theta}_j \tag{4.2-3}$$

where $\underline{L}_{j+1,j}$, not necessarily linear, may be assumed to be a state transition matrix (if $\underline{\theta}_j$ is treated as the state of the system at t_j) with the properties

$$\underline{L}_{j,j} = \underline{I} = \text{unity matrix for all } j$$

$$\underline{L}_{k,j} \underline{L}_{j,i} = \underline{L}_{k,i}$$

and

$$\underline{L}_{k,j}^{-1} = \underline{L}_{j,k}$$

Thus, if the law governing the variation of the optimum set is completely known, we can always take the time-varying effect into account and adjust the parameters

systematically to their optimum state. However, in most practical cases such as sonar detection problems, this variation is random due to the random nature of the unpredictable environments like thermal noise, surface agitation flow noise, cavitation noise, moving interferences, etc. Then the time-varying trend can only be described statistically by its measured or estimated frequency response or spectrum.

It seems likely that the well-known Wiener prediction theory can be applied to estimate the variation. But there is a serious drawback in using this theory. Since $\underline{\theta}_{j+1}$ is estimated from $\underline{\theta}_j$ and possibly other previous states, the problem at hand is similar to that of a random walk. The estimation error at each step may be small, but the accumulated error can be (not necessarily) very large. Convergence in mean square or in probability is not assured. In the next section we shall modify the sequency γ_j such that $\gamma_j' = \gamma_j + \beta_j$ in using the ordinary stochastic approximation method. The sequence γ_j satisfies the usual conditions $\gamma_j = \frac{\gamma}{j^\alpha}$, $\frac{1}{2} < \alpha \leq 1$ and β_j is used to correct the time-varying effect. As a limit $\gamma_j \to 0$ when $j \to \infty$, but β_j will converge to a small constant. Since the optimum set is always moving, some adjustments should be made at all times.

4.3 Application of the Kalman Filtering Techniques

It has been shown in our previous developments that the adaptive tapped-delay line filters designed via the methods of stochastic approximation using the current input information can asymptotically converge to the Wiener filters. Since Wiener filters are designed for stationary processes and their extensions to the time-varying case are the Kalman filters, we shall apply the Kalman filtering techniques to the design of adaptive tapped-delay-line filters with the hope that more rapidly convergent algorithms can be obtained and at the same time adaptation in nonstationary environments can be achieved. Consider a discrete filter consisting of tapped-delay-lines and K(M + 1) adjustable gains $\underline{W}$. Let Δ be the tap spacing

and C_{ik} be the weight at the k^{th} tap on the i^{th} filter.

Referring to Fig. 4 and the notation of Section 2.3, the filter output is given by

$$z(t) = \underline{W}^T \underline{n} \tag{4.3-1}$$

where $\underline{W}$ and $\underline{n}$ are K(M + 1)-dimensional vectors.

The performance criterion to be minimized is the mean squared error at each stage between the desired filter output $d(t)$ and the actual filter output $z(t)$.

$$J_j = \frac{1}{j} \sum_{k=1}^{j} [d_k - \underline{n}_k^T \underline{W}]^2 \triangleq \frac{1}{j} ||\underline{D}_j - \underline{H}_j \underline{W}||^2 \tag{4.3-2}$$

where D_j is a j-dimensional vector and $\underline{H}_j$ is a $j\times[(M+1)K]$-dimensional matrix defined respectively by

$$\underline{D}_j = \begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_j \end{bmatrix}, \quad \underline{H}_j = \begin{bmatrix} \underline{n}_1^T \\ \underline{n}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_j^T \end{bmatrix} \tag{4.3-3}$$

For large j and stationary environment Eq. (4.3-2) is just the usual mean squared error

$$J = \overline{e^2} = \overline{[d(t) - z(t)]^2} \tag{4.3-4}$$

Suppose that at each time t_j there are available the desired filter output D_j which is related to the optimum gains $\underline{\theta}_j$ and additive measuring noise $\underline{v}_j$

$$\underline{D}_j = \underline{H}_j \, \underline{\theta}_j + \underline{v}_j \tag{4.3-5}$$

The noise $\underline{v}_j$ is an additive random sequence with known statistics

$$E[\underline{v}_j] = 0 \quad \text{for all } j \tag{4.3-6}$$

$$E[\underline{v}_j \, \underline{v}_k^T] = \underline{\Phi}_j \, \delta_{jk} \tag{4.3-7}$$

The matrix $\underline{\phi}_j$ is assumed to be non-negative-definite unless otherwise stated. Let the optimum gains be generated by a source described by a first order dynamic system[1]

$$\underline{\theta}_j = a\,\underline{\theta}_{j-1} + \underline{u}_{j-1} \tag{4.3-8}$$

where the constant $0 < a < 1$ is a scalar and $\underline{u}_{j-1}$ is a vector random sequence with known statistics

$$E[\underline{u}_j] = 0 \tag{4.3-9}$$

$$E[\underline{u}_j\,\underline{u}_k^T] = \underline{Q}_k\,\delta_{jk} \tag{4.3-10}$$

The matrix $\underline{Q}_j$ is assumed to be nonnegative-definite, so it is possible that $\underline{u}_j = 0$. It is also assumed that the random sequences $\underline{v}_j$ and $\underline{u}_j$ are uncorrelated.

Employing the well-known Kalman filtering techniques (a summary is given in Appendix F), the estimate of the optimum gains at stage $(j + 1)$ can be calculated from its previous value by

$$\underline{W}_{j+1} = \underline{W}_j + \underline{K}_j\,(d_j - a\,\underline{n}_j^T\,\underline{W}_j \tag{4.3-11a}$$

$$\underline{K}_j = \underline{P}_j\,\underline{n}_j\,\underline{\phi}_j^{-1} \tag{4.3-11b}$$

$$\underline{P}_j^{-1} = (a^2\,\underline{P}_{j-1} + Q_{j-1})^{-1} + \underline{n}_j\,\underline{\phi}_j^{-1}\,\underline{n}_j^T \tag{4.3-11c}$$

---

1 It is to be noted that Eq. (4.3-8) is a rather particular model for which the results presented in this section hold true. This assumption makes the present approach a <u>partial</u> solution. For slowly varying parameters, the variance of the random sequence $\underline{u}_j$, var $[\underline{u}]$ = $(1-a^2)$ var $[\underline{\theta}]$, is small and permits us to assume $a \approx 1$.

(4.3-11c) can be rewritten in a more convenient form for computational purpose

$$\underline{P}_j = a^2\, \underline{P}_{j-1} + \underline{Q}_{j-1} - (a^2\, \underline{P}_{j-1} + \underline{Q}_{j-1})\, \underline{n}_j [\underline{n}_j^T (a^2 \underline{P}_{j-1} + \underline{Q}_{j-1}) \underline{n}_j$$

$$+ \ \phi_j]^{-1}\, \underline{n}_j^T\, (a^2 \underline{P}_{j-1} + \underline{Q}_{j-1}) \qquad (4.3\text{-}11d)$$

$\underline{P}$ is a K(M+1) x K(M+1)-dimensional matrix.

Since we are at liberty to process new data only one at a time, and for slowly time-varying case $a \approx 1$ [37], we arrive at the following simpler formulation for the iteration process

$$\underline{W}_{j+1} = \underline{W}_j + \underline{\Gamma}_{j+1}\, \underline{n}_{j+1} (d_{j+1} - \underline{n}_{j+1}^T\, \underline{W}_j) \qquad (4.3\text{-}12a)$$

with

$$\underline{\Gamma}_{j+1} = \frac{1}{\phi}\, \underline{P}_{j+1} \qquad (4.3\text{-}12b)$$

$$\underline{P}_{j+1}^{-1} = (\underline{P}_j + q\underline{I})^{-1} + \frac{1}{\phi}\, \underline{n}_{j+1}\, \underline{n}_{j+1}^T \qquad (4.3\text{-}12c)$$

or

$$\underline{P}_{j+1} = \underline{P}_j + q\underline{I} - \frac{(\underline{P}_j + q\underline{I})\underline{n}_{j+1}\, \underline{n}_{j+1}^T (\underline{P}_j + q\underline{I})}{[\ \phi + \underline{n}_{j+1}^T (\underline{P}_j + q\underline{I})\ \underline{n}_{j+1}\ ]} \qquad (4.3\text{-}12c)$$

where

$$\phi = \mathrm{Var}\ (v_j) \qquad (4.3\text{-}13)$$

$$q = \mathrm{Var}\ (u_j) \qquad (4.3\text{-}14)$$

for stationary white random sequences $\{v_j\}$ and $\{u_j\}$. Algorithm (4.3-12) will be discussed for both stationary and nonstationary cases. Its relationship to the method of stochastic approximation will also be given.

a) Stationary Case

In the stationary case the optimum gains are time invariant so that

$\underline{\theta}_j = \underline{\theta}_{j-1}$ for all j, $a = 1$ and $\underline{u}_j = 0$. Suppose that $\phi = 1$ *, (3.6-15) reduces to

$$\underline{W}_{j+1} = \underline{W}_j + \underline{\Gamma}_{j+1}\, \underline{n}_{j+1}\, (d_{j+1} - \underline{n}_{j+1}^T\, \underline{W}_j) \tag{4.3-15a}$$

$$\underline{\Gamma}_{j+1}^{-1} = \Gamma_j^{-1} + \underline{n}_{j+1}\, \underline{n}_{j+1}^T \tag{4.3-15b}$$

Combining Eq. (4.3-15a) and (4.3-15b) gives an alternative expression

$$\underline{W}_{j+1} = [\underline{H}_j^T\, \underline{H}_j]^{-1}\, \underline{H}_j^T\, \underline{D}_j \tag{4.3-16}$$

which is just the solution of minimizing the error (4.3-4) by the least square fit. The relationship between optimal filtering and least square fit has been pointed out by various authors [43, 44]. The sequence $\underline{W}_j$ described by Eq. (4.3-16) will be shown to converge to the optimum gains

$$\underline{\theta} = \underline{R}_n^{-1}\, \underline{R}_{dn} = \underline{R}_n^{-1}\, \underline{R}_{d\xi} \tag{4.3-17}$$

This is done by rewriting Eq. (4.3-16) in the form

$$\underline{W}_{j+1} = [\, \frac{1}{j}\, \underline{H}_j^T\, \underline{H}_j]^{-1}\, [\, \frac{1}{j}\, \underline{H}_j^T\,]\underline{D}_j \tag{4.3-18}$$

Applying the strong law of large numbers and making use of the fact that continuous functions of convergent random sequences are also convergent, we can state that

$$\lim_{j\to\infty} [\frac{1}{j}\, \underline{H}_j^T\, \underline{H}_j]^{-1} = \lim_{j\to\infty} [\frac{1}{j} \sum_{k=1}^{j} \underline{n}_k\, \underline{n}_k^T]^{-1}$$

$$= \{E_x\, [\underline{n}\, \underline{n}^T]\}^{-1} = \underline{R}_n^{-1} \tag{4.3-19}$$

---

* This corresponds to the problem of minimizing $||\underline{D}_j - \underline{H}_j\, \underline{\theta}||^2$ instead of $||\underline{D}_j - \underline{H}_j\, \underline{\theta}||^2_{\underline{R}_j^{-1}}$ used in Eq. (F-7). The weighting matrix $\underline{R}_j^{-1}$ is required is required if the measuring noise $\underline{v}_j$ is Gaussian distributed with zero mean and covariance matrix $\underline{R}_j$. For details, see [44].

with probability one if $\underline{R}_\eta$ exists and is positive definite. $\underline{R}_\eta$ always meets these requirements because it is the correlation matrix of the delayed inputs. By the same token we can state

$$\lim_{j\to\infty} [\frac{1}{j} \underline{H}_j^T \underline{D}_j] = \lim_{j\to\infty} [\frac{1}{j} \sum_{k=1}^{j} d_k \underline{\eta}_k] = \underline{R}_{d\eta} \tag{4.3-20}$$

with probability one if $R_{d\eta}$ exists. Consequently one concludes that

$$\text{Prob} \left\{ \lim_{j\to\infty} (\underline{W}_j) = \underline{R}_\eta^{-1} \underline{R}_{d\eta} = \underline{\theta} \right\} = 1 \tag{4.3-21}$$

and the limit $\underline{\theta}$ minimizes the mean squared error defined by Eq. (4.3-2). It is noted that in using algorithm (4.3-15) the initial estimate of the parameters $\underline{W}_1$ is arbitrary and Γ_o^{-1} is finite and positive definite. We can just set $\underline{W}_1 = 0$ and $\Gamma_o = \underline{I}$ = unit matrix to start the iteration process. The connection between algorithm (4.3-15) and the ordinary method of stochastic approximation can be constructed as follows:

From Eq. (4.3-15b) we have

$$\begin{aligned} \underline{\Gamma}_j^{-1} &= \underline{\Gamma}_{j-1}^{-1} + \underline{\eta}_j \underline{\eta}_j^T = \underline{\Gamma}_o^{-1} \sum_{k=1}^{j} \underline{\eta}_k \underline{\eta}_k^T \\ &= \underline{\Gamma}_o^{-1} + j(\frac{1}{j} \sum_{k=1}^{j} \underline{\eta}_k \underline{\eta}_k^T) \end{aligned} \tag{4.3-22}$$

which for large j converges to

$$\lim_{j\to\infty} \underline{\Gamma}_j^{-1} = \underline{\Gamma}_o^{-1} + j \underline{R}_\eta \simeq j \underline{R}_\eta \tag{4.3-23}$$

Thus the weighting matrix appearing in Eq. (4.3-15) approaches to

$$\lim_{j\to\infty} \Gamma_j \simeq \frac{1}{j} \underline{R}_\eta^{-1}$$

and the corresponding algorithm becomes

$$\underline{W}_{j+1} = \underline{W}_j + \frac{1}{j+1} \underline{R}_\eta^{-1} \underline{\eta}_{j+1}(d_{j+1} - \underline{\eta}_{j+1}^T \underline{W}_j) \tag{4.3-24}$$

Eq. (4.3-24) is just the adjustment procedure derived from the ordinary method of stochastic approximation

$$\underline{W}_{j+1} = \underline{W}_j + \underline{\gamma}_{j+1}\, \underline{\eta}_{j+1} (d_{j+1} - \underline{\eta}^T_{j+1} \underline{W}_j) \tag{4.3-25}$$

with

$$\underline{\gamma}_{j+1} = \frac{1}{j+1}\, \underline{R}_\eta^{-1} \tag{4.3-26}$$

$\underline{\gamma}_{j+1}$ is considered as a weighting matrix in this case and in the simplest case is just $\underline{\gamma}_{j+1} = \frac{\gamma}{j+1}\underline{I}$, if $\underline{R}_\eta$ is a diagonal matrix and $\gamma =$ constant. In the transformed parameter space where $\underline{W}' = \underline{P}\,\underline{W}$, $R_\eta = \underline{P}^{-1}\underline{\Lambda}\,\underline{P}$ the optimum weighting sequence is $\gamma^{(k)}_{j+1} = \frac{1}{(j+1)\lambda_k}$, λ_k being the k^{th} eigenvalue of $\underline{R}_\eta$. Explicit expressions for the rate of convergence have been derived in Section 3.5. Since algorithms (4.3-15) and (4.3-25) minimize the same quadratic criterion in the limit, it will be of interest to compare the convergence properties of the two algorithms. It is to be expected that algorithm (4.3-15) will be more rapidly convergent than algorithm (4.3-25) since at each stage of the iteration, algorithm (4.3-15) uses information from the inputs of all past stages whereas algorithm (4.3-25) only uses information from the input that is received at the current stage. The optimum sequence $\{\gamma_j\}$ defined by Eq. (4.3-26) can be predetermined only when we know the correlation matrix $\underline{P}_\eta$, which contradicts our motivation of using adaptive techniques. That algorithm (4.3-15) converges faster than algorithm (4.3-25) with arbitrary $\gamma_j = \frac{\gamma}{j^\alpha}$, $\frac{1}{2} < \alpha \leq 1$, can be further illustrated by comparing the methods of minimization embodied in each of the algorithms. Algorithm (4.3-25) proceeds in the direction of the negative gradient of $(d_j - \underline{W}^T \underline{\eta}_j)^2$ at the j^{th} iteration stage. Algorithm (4.3-15), on the other hand, is a second order algorithm that selects $\underline{W}_j$ which minimizes $\sum_{k=1}^{j}[(d_k - \underline{W}^T \underline{\eta}_k)^2]$ at the j^{th} iteration stage. On this basis, it seems plausible that algorithm (4.3-15) should be more rapidly convergent than algorithm (4.3-25), a conjecture

that will be reinforced by the simulation results presented in Chapter Six. The requirement that the desired filter output $d(t)$ be available can be removed using signal correlation functions.

(b) Nonstationery Case

In the general case where the optimum gains vary with time as a result of the nonstationary noise fields, algorithm (4.3-12) will have to be used to make adaptation possible in nonstationary environments. If we are willing to adjust the weighting matrix $\underline{\Gamma}_{j+1}$ at each stage during the adaptation period, algorithm (4.3-12) is the desired procedure. If, however, we just want to modify the ordinary method of stochastic approximation such that

$$\gamma_j' = \gamma_j + \beta \tag{4.3-27}$$

where

$$\gamma_j = \frac{\gamma}{j^{\alpha}} \quad , \quad \frac{1}{2} < \alpha \leq 1$$

then β = constant can be found in the following discussion to correct the time-varying effect. As a limit $\gamma_j \to 0$ when $j \to \infty$, but adjustments are made at all time due to the presence of β.

Consider the adjustment procedure

$$\underline{W}_{j+1} = \underline{W}_j + \underline{\Gamma}_{j+1} \underline{n}_{j+1} \sigma^{-1} (d_{j+1} - \underline{W}_j^T \underline{n}_{j+1}) \tag{4.3-28}$$

where the weighting matrix $\underline{\Gamma}$ in the stationary case is

$$\underline{\Gamma}_{j+1}^{-1} = \underline{\Gamma}_j^{-1} + \frac{1}{\sigma} \underline{n}_{j+1} \underline{n}_{j+1}^T \tag{4.3-29}$$

and that in the nonstationary case is

$$\underline{\Gamma}_{j+1}'^{-1} = (\underline{\Gamma}_{j+1} + \underline{B}_{j+1})^{-1} = (\underline{\Gamma}_j + \underline{B}_j + q\,\underline{I})^{-1} + \frac{1}{\sigma} \underline{n}_{j+1} \underline{n}_{j+1}^T \tag{4.3-30}$$

Following the arguments leading to Eq. (4.3-23), we can write from Eq. (4.3-29)

$$\underline{\Gamma}_{j+1}^{-1} \underset{j \to \infty}{\approx} \frac{(j+1)}{\sigma} \underline{R} \tag{4.3-31}$$

Since Eq. (4.3-30) is a nonlinear difference equation of the form

$$x_{j+1}^{-1} = (x_j + a)^{-1} + y_{j+1}$$

whose explicit solution is not available, we shall only consider the asymptotical behavior of Eq. (4.3-30) for large j such that

a. $\underline{\Gamma}_{j+1}^{-1} \simeq \underline{\Gamma}_j^{-1}$

b. $\frac{1}{\phi}\ \underline{n}_{j+1}\ \underline{n}_{j+1}^T << [\underline{\Gamma}_{j+1} + \underline{B}_{j+1}]^{-1}$

c. $\underline{\Gamma}_{j+1} << \underline{B}_{j+1}$

If the above assumptions hold, one can obtain the steady state $\underline{B}$ as*

$$\lim_{j\to\infty} \underline{B}_j = q \qquad (4.3\text{-}32)$$

and the adjustment constant β defined in Eq. (4.3-27) has the form

$$\beta = \frac{q}{\phi} \qquad (4.3\text{-}33)$$

Therefore, a simple iterative procedure to adjust the gains in nonstationary environment is

$$\underline{W}_{j+1} = \underline{W}_j + (\gamma_j + \frac{q}{\phi})\ \underline{n}_{j+1}\ (d_{j+1} - \underline{W}_j^T\ \underline{n}_{j+1}) \qquad (4.3\text{-}34)$$

The quantity $d_{j+1}\ \underline{n}_{j+1}$ appearing in Eq. (4 3-34) can be replaced by the signal correlation function $\underline{R}_{d\xi}$.

Summary - Let the actual target wavefront d_j and the summer output $z_j = \underline{n}_j^T\ \underline{W}_j$ be related at time t_j by

$$d_j = \underline{n}_j^T\ \underline{W}_j + v_j \qquad (4.3\text{-}35)$$

where the noise v_j is an additive random sequence with known statistics

$$E[v_j] = 0\ ,\ \text{Var}\ (v_j) = \phi\ .$$

* See Appendix G for details

Let the optimum gains $\underline{\hat{}}_j$ be described by a first order dynamic system

$$\underline{\hat{}}_j = a\, \underline{\hat{}}_{j-1} + \underline{u}_{j-1} \tag{4.3-36}$$

where the constant $0 < a < 1$ is a scalar and $\underline{u}_j$ a vector random sequence with known statistics $E[\underline{u}_j] = 0$, $E[\underline{u}_j\, \underline{u}_j^T] = q\, I$. The algorithm to estimate the adjustable gains $\underline{W}$ which minimizes the mean squared error at each stage

$$J_j = \frac{1}{j} \sum_{k=1}^{j} [d_k - \underline{n}_k^T\, \underline{W}] \tag{4.3-37}$$

can be written in the general form

$$\underline{W}_{j+1} = \underline{W}_j + \Gamma_j \quad [\underline{R}_{d\xi} - \underline{n}_j\, \underline{n}_j^T\, \underline{W}_j] \tag{4.3-38}$$

The weighting sequence $\underline{\Gamma}_j$ assumes different expressions depending on the stationarity of process and on whether an optimum estimation procedure (fastest convergence rate) is required. Four different choices of $\underline{\Gamma}_{j+1}$ are listed below:

stationarity / optimality	stationary	nonstationary
optimum	$\underline{\Gamma}_j = \underline{\Gamma}_{j-1}^{-1} + \underline{n}_j\, \underline{n}_j^T\, \phi^{-1}$	$\underline{\Gamma}_j^{-1} = (\underline{\Gamma}_{j-1} + q\underline{I})^{-1} + \frac{1}{\phi}\, \underline{n}_j\, \underline{n}_j^T$
nonoptimum	$\underline{\Gamma}_j = \gamma_j\, \underline{I}$	$\underline{\Gamma}_j = (\gamma_j + \frac{q}{\phi})\, \underline{I}$

where $\gamma_j = \frac{\gamma}{j^\alpha}$, $\gamma > 0$, $\frac{1}{2} < \alpha < 1$.

4.4 Nonstationarity and the use of Ordinary Methods of Stochastic Approximation

a) Notations

Recall that the adjustment procedure used very frequently in this study is of the form

$$\underline{W}_{j+1} = \underline{W}_j + 2\gamma_j\, \underline{R}_{d\xi} - 2\gamma_j\, z_j\, \underline{n}_j \tag{4.4-1}$$

which is derived from the ordinary method of stochastic approximation

$$\underline{W}_{j+1} = \underline{W}_j - \gamma_j \nabla Q \tag{4.4-2}$$

The optimum gains for the multiple-sensor array

$$\underline{\theta} = \underline{W}_{op} = (\underline{R}_\xi + \underline{R}_\nu)^{-1} \underline{R}_{d\xi} \tag{4.4-3}$$

are obtained by solving $E[\nabla Q] = 0$.

In order to illustrate the essential steps involved, we shall consider the simplest case where only a single filter is designed by adjusting a single gain constant. Extensions to the general case is reasonably straightforward.

The optimum gain in this case is then

$$\theta = c_{op} = [R_s(0) + R_n(0)]^{-1} R_s(0) \tag{4.4-4}$$

where $R_s(\tau)$ and $R_n(\tau)$ are respectively the signal and noise autocorrelation functions. When the noise field is time-varying such that

$$R_n(\tau,t) = R_n(\tau)\ [1 + f(t)] \tag{4.4-5}$$

the optimum gain is also a function of time t

$$\theta(t) = R_s(0) + R_n(0)\ [1 + f(t)]^{-1} R_s(0) \tag{4.4-6}$$

whose value at time $t = j$ will be denoted by θ_j .

In the above $f(t)$ is a time function and depends entirely on the nonstationarity of noise fields. This function will be a constant in the stationary case.

At time $t = j$, the optimum gain can be written from Eq. (4.4-6) as

$$\theta_j = \theta_o + F_j \tag{4.4-7}$$

where

$$\theta_o = [R_s(0) + R_n(0)]^{-1} R_s(0) \tag{4.4-8}$$

is the time-invariant optimum gain and

$$F_j = -\theta_o \left\{ [R_s(0) + R_n(0)]^{-1} R_n(0)f + [R_s(0) + R_n(0)]^{-2} R_n^2(0)f_j^2 + \cdots \right\} \tag{4.4-9}$$

is the time-varying part resulting from a simple series expansion.

Thus, the optimum gain at any instant can be related to its previous value by

$$\theta_{j+1} = \theta_j + \Delta\theta_j \tag{4.4-10}$$

$\Delta\theta_j$ is the increment of θ_j at $t = j$ and given by

$$\begin{aligned} \Delta\theta_j &= F_{j+1} - F_j \\ &= \theta_o(f_{j+1} - f_j)R_n(0) \left\{ -1 + [R_s(0) + R_n(0)]^{-1}R_n(0)(f_{j+1} + f_j) + \cdots \right\} \end{aligned} \tag{4.4-11}$$

Let the dominating part of $\Delta\theta_j$ be denoted by $0(\frac{1}{j^\omega})$.

Then we have

$$\theta_{j+1} = \theta_j + 0(\frac{1}{j^\omega}) \tag{4.4-12}$$

Up to now we have not posed any restrictions on ω so that Eq. (4.4-12) is valid in general.

b) <u>Assumptions and Analysis</u>

It is assumed that in using the adjustment procedure

$$c_{j+1} = c_j - \gamma_j \nabla Q_j \tag{4.4-13}$$

the following conditions hold

(i) There exist constants K_ℓ and K_u such that

$$K_\ell | c_j - \theta_j | < |\overline{\nabla Q_j}| < K_u | c_j - \theta_j | \tag{4.4-14}$$

for all j . This simply says that the gradient is of bounded variation.

Note that

$$\overline{\nabla Q}_j = \overline{\nabla Q}(c_j) \lesssim 0 \quad \text{when} \quad c_j \gtrless \theta_j \tag{4.4-15}$$

and

$$E[\nabla Q_j | c_1 \cdots, c_j] = (c_j - \theta_j)\mu \tag{4.4-16}$$

for some μ , $K_\ell < \mu < K_u$.

(ii) The conditional variance of ∇Q_j is also bounded

$$\text{Var}\,[\nabla Q_j | c_1 \cdots, c_j] \leq \sigma^2 < \infty \tag{4.4-17}$$

Since $\text{Var}\,[x] = \overline{x^2} - \bar{x}^2$, we can write from Eqs. (4.4-16) and (4.4-17)

$$\begin{aligned} &E[(\nabla Q_j)^2 | c_1, \cdots, c_j] \\ &= \text{Var}\,[\nabla Q_j | c_1, \cdots, c_j] + \left\{ E[\nabla Q_j | c_1, \cdots, c_j] \right\}^2 \\ &\leq \sigma^2 + (c_j - \theta_j)^2 \mu^2 \end{aligned} \tag{4.4-18}$$

(iii) The weighting sequence is of the form

$$\gamma_j = \frac{\gamma}{j^\alpha} \; , \quad \gamma > 0 \; , \; \frac{1}{2} < \alpha \leq 1 \; . \tag{4.4-19}$$

Subtracting Eq. (4.4-13) by Eq. (4.4-12)

$$c_{j+1} - \theta_{j+1} = c_j - \theta_j - \gamma_j \nabla Q_j - 0 \left(\frac{1}{j^\omega}\right) \tag{4.4-20}$$

and squaring give

$$\begin{aligned} (c_{j+1} - \theta_{j+1})^2 = c_j - \theta_j)^2 + \gamma_j^2 (\nabla Q_j)^2 + 0 \left(\frac{1}{j^{2\omega}}\right) \\ - 2\gamma_j \nabla Q_j (c_j - \theta_j) + 2\gamma_j \nabla Q\, 0\left(\frac{1}{j^\omega}\right) \\ - 2 (c_j - \theta_j)\, 0\left(\frac{1}{j^\omega}\right) \end{aligned} \tag{4.4-21}$$

Taking the conditional expectations on both sides of Eq. (4.4-21) yields

$$E[(c_{j+1} - \theta_{j+1})^2 | c_1, \cdots, c_j]$$

$$= (c_j - \theta_j)^2 + \gamma_j^2 E[(\nabla Q_j)^2 | c_1, \cdots c_j] + \frac{K_1}{j^{2\omega}}$$

$$- 2 \gamma_j (c_j - \theta_j) E[\nabla Q_j | c_1, \cdots, c_j]$$

$$+ 2 \gamma_j O(\frac{1}{j^\omega}) E[\nabla Q_j | c_1, \cdots, c_j]$$

$$- 2 (c_j - \theta_j) O(\frac{1}{j^\omega}) \qquad (4.4\text{-}22)$$

From Eqs. (4.4-16) and (4.4-18) it follows that

$$E[(c_{j+1} - \theta_{j+1})^2 | c_1, \cdots, c_j]$$

$$\leq (c_j - \theta_j)^2 + \gamma_j^2 [\sigma^2 + (c_j - \theta_j)^2 \mu^2 + \frac{K_1}{j^{2\omega}}$$

$$- 2 \gamma_j (c_j - \theta_j)^2 \mu + 2 \gamma_j O(\frac{1}{j^\omega}) | c_j - \theta_j | \mu$$

$$+ 2 \gamma_j O(\frac{1}{j^\omega}) | c_j - \theta_j |$$

$$= (c_j - \theta_j)^2 + \frac{K_2}{j^{2\alpha}} + \frac{K_3}{j^{2\alpha}} (c_j - \theta_j)^2 + \frac{K_1}{j^{2\omega}}$$

$$- \frac{K_4}{j^\alpha} (c_j - \theta_j)^2 + \frac{K_5}{j^\alpha} \frac{1}{j^\omega} | c_j - \theta_j |$$

$$+ \frac{K_6}{j^\omega} | c_j - \theta_j | \qquad (4.4\text{-}23)$$

where the K_i's are constants.

We shall now consider several ranges of ω relative to α .

Case 1. $\omega > \alpha$.

After enlarging the corresponding coefficients, the terms of lower order of

magnitude will include the terms of higher order. Thus, for $\omega > \alpha$,

Eq. (4.4-23) becomes

$$E[(c_{j+1} - \theta_{j+1})^2 | c_1, \ldots, c_j]$$

$$\leq (1 - \frac{K_4}{j^\alpha})(c_j - \theta_j)^2 + \frac{K_2}{j^{2\alpha}} + \frac{K_6}{j^\omega} | c_j - \theta_j | \qquad (4.4\text{-}24)$$

Now, we take (unconditional) expectations on both sides of Eq. (4.4-24). When estimating $E[| c_j - \theta_j |]$, we use the inequality [47]

$$E[|x|] < \varepsilon + \varepsilon^{-1} E[x^2] \qquad (4.4\text{-}25)$$

The inequality (4.4-25) holds true for every $\varepsilon > 0$ and every random variable with finite variance.

If we set $\varepsilon = \frac{1}{\delta j^{\omega-\alpha}}$ for some small $\delta > 0$, then the unconditional expectation of Eq. (4.4-24) is

$$E[c_{j+1} - \theta_{j+1})^2]$$

$$\leq (1 - \frac{K_4}{j^\alpha}) E[(c_j - \theta_j)^2] + \frac{K_2}{j^{2\alpha}}$$

$$+ \frac{K_6}{j^\omega} \left\{ \frac{1}{\delta j^{\omega-\alpha}} + \delta j^{\omega-\alpha} E[(c_j - \theta_j)^2] \right\}$$

$$= (1 - \frac{K_7}{j^\alpha}) E[(c_j - \theta_j)^2] + \frac{K_1}{j^{2\omega-\alpha}} + \frac{K_2}{j^{2\alpha}} \qquad (4.4\text{-}26)$$

A lemma due to K. L. Chung* will be used here.

<u>Chung's lemma</u> : Let V_j , $j = 1,2,\ldots\ldots\ldots$, be real numbers such that for $j > J$

$$V_{j+1} < (1 - \frac{a}{j^s}) V_j + \frac{b}{j^t} \qquad (4.4\text{-}27)$$

---

* Chung, K. L., "On a Stochastic Approximation Method," <u>Ann. Math. Statist</u>. vol. 25, pp. 463-483, 1954.

where $0 < s < 1$, $a > 0$, $b > 0$, t real.

Then

$$\lim_{n\to\infty} \sup j^{t-s} V_j \leq \frac{b}{a} \tag{4.4-28}$$

This lemma remains true if the inequalities (4.4-27) and (4.4-28) are reversed and simultaneously, $\lim\limits_{n\to\infty} \sup$ is changed into $\lim\limits_{n\to\infty} \inf$.

Upon using Eqs. (4.4-27) and (4.4-28), we have

$$E[(c_j - \theta_j)^2] \approx O(j^{-\alpha}) \quad \text{for} \quad \omega \geq \frac{3}{2}\alpha$$

and

$$E[(c_j - \theta_j)^2 \approx O(j^{-2\omega + 2\alpha}) \quad \text{for} \quad \alpha < \omega < \frac{3}{2}\alpha \tag{4.4-29}$$

<u>Case 2</u>. $\omega < \alpha$.

Under this situation Eq. (4.4-23) reduces to

$$E[c_{j+1} - \theta_{j+1})^2 | c_1, \cdots c_j]$$

$$\leq (1 - \frac{K_4}{j^\alpha})(c_j - \theta_j)^2 + \frac{K_6}{j^\omega}|c_j - \theta_j| + \frac{K_1}{j^{2\omega}} \tag{4.4-30}$$

If we take the unconditional expectations on both sides of Eq. (4.4-30) and follow similar steps leading from Eq. (4.4-24) to Eq. (4.4-26), we obtain

$$E[(c_{j+1} - \theta_{j+1})^2] \leq (1 - \frac{K_4}{j^\alpha}) E[(c_j - \theta_j)^2] + \frac{K_7}{j^{2\omega-\alpha}} \tag{4.4-31}$$

Invoking Chung's lemma gives

$$E[(c_j - \theta_j)^2] \approx O(j^{-2\alpha-2\omega}) \tag{4.4-32}$$

Since, by assumption in this case $\omega < \alpha$, we see that the sequence $\{c_j\}$, $j = 1,2,\ldots\ldots$, will diverge.

<u>Case 3.</u> $\omega = \alpha$.

Following steps similar to the above two cases and letting ε be a constant,

we have

$$E[(c_{j+1} - \theta_{j+1})^2 \mid c_1, \cdots, c_j] \leq (1 - \frac{K_4}{j^\alpha})(c_j - \theta_j)^2 + \frac{K_1}{j^{2\alpha}}$$

$$+ \frac{K_6}{j^\alpha} |c_j - \theta_j| \qquad (4.4\text{-}33)$$

and

$$E[(c_{j+1} - \theta_{j+1})^2] \leq (1 - \frac{K_4}{j^\alpha}) E[(c_j - \theta_j)^2] + \frac{K_1}{j^{2\alpha}}$$

$$+ \frac{K_6}{j^\alpha} \quad \varepsilon + \varepsilon^{-1} E[(c_j - \theta_j)^2]$$

$$= (1 - K_5/j^\alpha) E[(c_j - \theta_j)^2] + K_6/j^\alpha$$

$$(4.4\text{-}34)$$

Chung's lemma gives

$$E[(c_j - \theta_j)^2] \approx O(j^0) = \text{constant} \qquad (4.4\text{-}35)$$

c) Conclusion

There are several points worth noting in the above analysis. All the results are intuitively reasonable. The inequality $\omega > \alpha$ indicates that the rate of parameter variation is slower than the rate of convergence in the stationary case (in the order of $j^{-\alpha}$). If the rate of time-variation is relatively slow ($\omega \geqslant \frac{3}{2}\alpha$), the ordinary method of stochastic approximation can be employed to adjust the time-varying parameter without affecting the rate of convergence. On the other hand, if the optimum parameters vary at a rate slower than but comparable to the rate at which γ_j decreases ($\alpha < \omega < \frac{3}{2}\alpha$), the actual rate of convergence is reduced by an amount depending upon the difference $(\omega-\alpha)$. Suppose that the rate of parameter variation is faster than that of convergence, we can never expect to have the algorithm converge to the desired value at any time.

This is indicated in Case 2 where $\omega < \alpha$. In short, the interrelationships of ω, α, and the rate of convergence are

$$\omega < \alpha \qquad V_j \triangleq \overline{(c_j - \theta_j)^2} \longrightarrow \infty$$

$$\omega = \alpha \qquad V_j \longrightarrow \text{constant}$$

$$\alpha < \omega < \frac{3}{2}\alpha \qquad V_j = O(j^{-2\omega-2\alpha})$$

$$\omega \geq \frac{3}{2}\alpha \qquad V_j = O(j^{-\alpha})$$

CHAPTER FIVE

PERFORMANCE ANALYSIS OF THE ADAPTIVE RECEIVER

5.1 Introduction and Assumptions

In Chapter Two the rather practical situation in which filters in an array consist of weighted-tapped-delay lines were considered. It is shown that tapped-delay-line filters can approximate the continuous Wiener filters quite closely with proper delays and proper weights. The method of stochastic approximation and a mean square error criterion were employed in Chapter Three to derive adjustment procedures using signal statistics.

An adaptive array receiver is formed by incorporating these adaptive tapped-delay-line filters in an array as shown in Fig. 3. In this chapter we shall study the performance of such a processor. The performance criteria to be evaluated are the output signal-to-noise ratio and directivity patterns. These quantities depend on a number of system parameters such as field (target, noise, interferences) properties, number of hydrophones in an array, number of taps and their spacings on the tapped-delay lines, adaptation time, locations of the target and interferences, etc.

The following assumptions are used to simplify the analyses:

1) Target, interference and ambient noise are assumed to be Gaussian random processes.
2) The receiving array is assumed to be linear and to consist of K omnidirectional hydrophones.
3) The wavefronts of target signal and interference are regarded as plane over the dimensions of the receiving array.
4) The sum of interference and ambient noise is regarded as the effective noise.

5) The input spectra are identical in shapes (but not in levels) over the frequency range $(0, \omega_o)$ where most of their power is concentrated. This situation closely resembles conditions encountered in practice if one ignores periodic components of the input processes.

6) The noise consists of a single point interference and ambient noise.

7) The ambient noise is statistically independent from hydrophone to hydrophone.

Mathematically, the above assumptions are equivalent to the following equations.

(1) Ratios of the input

$$\begin{aligned} \phi_d(\omega)/\phi_n(\omega) &= S/N \\ \phi_n(\omega)/\phi_I(\omega) &= N/I \\ \phi_I(\omega)/\phi_d(\omega) &= I/S \end{aligned} \qquad (5.1\text{-}1)$$

for $0 \leq \omega \leq \omega_o$, ω_o is large. The spectra are zero elsewhere.

(2) Spectral Matrices

$$\begin{aligned} &\text{Signal} \quad \underline{\Phi}_{ss} = \phi_d \, \underline{a} \, \underline{a}^{*T} \\ &\text{Noise} \quad \underline{\Phi}_{nn} = \phi_n \, \underline{I} + \phi_I \, \underline{b} \, \underline{b}^{*T} \end{aligned} \qquad (5.1\text{-}2)$$

With the aid of a matrix inversion formula [7][1]

$$[\underline{p} + \underline{q}\underline{\gamma}^T]^{-1} = \underline{p}^{-1} - \{(\underline{p}^{-1}\underline{q})(\underline{\gamma}^T\underline{p}^{-1})/(1 + \underline{\gamma}^T \underline{p}^{-1}\underline{q})\}$$

we have

$$\underline{\Phi}_{nn}^{*-1} = \frac{1}{\phi_n} [\underline{I} - \frac{\underline{b}^* \underline{b}^T}{K + \frac{\phi_n}{\phi_I}}] \qquad (5.1\text{-}3)$$

---

[1] If $\underline{p}^{-1}$ exists and $\underline{q} \, \underline{\gamma}^T$ is of rank 1.

where

$$\underline{a}^T = [e^{j\omega\tau_1}\ e^{j\omega\tau_2} \ldots e^{j\omega\tau_K}]$$

$$\underline{b}^T = [e^{j\omega\rho_1}\ e^{j\omega\rho_2} \ldots e^{j\omega\rho_K}] \qquad (5.1\text{-}4)$$

and for a linear array of equally spaced hydrophones,

$$\tau_i - \tau_h = |i - h|\frac{d}{c} \sin\theta_T\ ,\ \rho_i - \rho_h = |i - h|\frac{d}{c} \sin\theta_I \qquad (5.1\text{-}5)$$

together with the following definitions

d = hydrophone spacing

c = sound velocity in water

θ_T = target angle

θ_I = interference angle

$\underline{I}$ = unity matrix

5.2 Statistics of an Array Receiver

It will be necessary for all cases to obtain expressions for the mean and variance of the detector output. Some useful expressions for the related spectral densities and spectral matrices will be obtained first. Referring to Figs. 1 to 3, the beamformer output is

$$z(t) = \sum_{i=1}^{K} \gamma_i(t) \qquad (5.2\text{-}1)$$

and, therefore, its autocorrelation function is

$$R_z(\tau) = E\{z(t)\ z(x+\tau)\} = \sum_{i=1}^{K}\sum_{k=1}^{K} E\{\gamma_i(t)\ \gamma_k(x+\tau)\}$$

$$= \sum_{i=1}^{K}\sum_{k=1}^{K} R_{\gamma_i\gamma_k}(\tau) \qquad (5.2\text{-}2)$$

The power spectral density of z is consequently

$$\phi_z(\omega) = \sum_{i=1}^{K}\sum_{k=1}^{K} \phi_{\gamma_i\gamma_k}(\omega) = \sum_{i=1}^{K}\sum_{k=1}^{K} \phi_{x_ix_k}(\omega)\ H_i(\omega)\ H_k^*(\omega) \qquad (5.2\text{-}3)$$

where * indicates the conjugate. If the transfer function vector $\underline{H}(\omega)$ is defined by

$$\underline{H}^T(\omega) = [H_1(\omega)\ H_2(\omega)\ \cdots\ H_K(\omega)] \tag{5.2-4}$$

then Eq. (5.2-3) can be written compactly as

$$\phi_z(\omega) = \underline{H}^T(\omega)\ \underline{\Phi}_{xx}(\omega)\ \underline{H}^*(\omega) \tag{5.2-5}$$

where $\underline{H}^T(\omega)$ is the transpose of $\underline{H}(\omega)$ and

$$\underline{\Phi}_{xx} = \begin{bmatrix} \phi_{x_1} & \cdots & \phi_{x_1x_k} \\ \vdots & & \vdots \\ \phi_{x_Kx_1} & & \phi_{x_Kx_K} \end{bmatrix} \tag{5.2-6}$$

is the input spectral matrix.

If the referenced signal $d(t)$ and noise $n_i(t)$ are assumed to be uncorrelated for all i, then

$$R_{x_ix_k}(\tau) = R_{n_in_k}(\tau) + R_{s_is_k}(\tau)$$

$$\phi_{x_ix_k}(\omega) = \phi_{n_in_k}(\omega) + \phi_{s_is_k}(\omega)$$

and

$$\underline{\Phi}_{xx} = \underline{\Phi}_{ss} + \underline{\Phi}_{nn} \tag{5.2-7}$$

The $\underline{\Phi}$'s are understood to be functions of ω. $\underline{\Phi}_{ss}$ and $\underline{\Phi}_{nn}$ are respectively the signal and noise spectral matrices defined by Eqs. (2.1-3). One very useful property of the spectral matrices is that they are Hermitian, i.e.,

$$\underline{\Phi}^{*T} = \underline{\Phi} \tag{5.2-8}$$

Assuming without loss of generality[1] that the averaging filter has unity gain at

---

[1] This assumes that the filter does not have any poles at $\omega=0$; i.e., does not contain an integrator. Thus this assumption is not completely general, but in practice integrators will always be of finite time. So it is not a very serious loss of generality.

$\omega = 0$, the detector output will have an average

$$\bar{y} = \int_0^\infty h_{av}(\tau)\, \overline{z_2(x-\tau)}\, d\tau = \bar{z}_2\, H_{av}(\omega = 0)$$

$$= \bar{z}_2 = \overline{z_1^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{z_1}(\omega)\, d\omega \tag{5.2-9}$$

But

$$\phi_{z_1}(\omega) = |G(\omega)|^2\, \phi_z(\omega) \tag{5.2-10}$$

Hence, in the presence of signal, using Eq. (5.2-5)

$$\langle y \rangle_{S+N} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{H}^T\, \underline{\Phi}_{xx}\, \underline{H}^*\, d\omega \tag{5.2-11}$$

in the absence of signal

$$\langle y \rangle_N = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{H}^T\, \underline{\Phi}_{nn}\, \underline{H}^*\, d\omega \tag{5.2-12}$$

and the d.c. change of the output becomes

$$\bar{y}_{d.c.} = \langle y \rangle_{S+N} - \langle y \rangle_N = \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{H}^T\, \underline{\Phi}_{ss}\, \underline{H}^*\, d\omega \tag{5.2-13}$$

In order to obtain a convenient expression for the output variance, assume that the averaging time T_{av} is long compared to the correlation time of $z_1(t)$, or equivalently, that the bandwidth of the lowpass averaging filter $H_{av}(\omega)$ is much narrower than the bandwidth of $z_1(t)$. Then [23, 24]

$$\sigma_y^2 \simeq \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} [\phi_{z_1}(\omega)]^2\, d\omega$$

$$= \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4\, \{\underline{H}^T(\underline{\Phi}_{ss} + \underline{\Phi}_{nn})\, \underline{H}^*\}^2\, d\omega \tag{5.2-14}$$

where T_{av} is the averaging time defined by

$$1/T_{av} \triangleq \int_{0}^{\infty} h_{av}^{2}(\tau)d\tau \tag{5.2-15}$$

So far we have assumed that the individual filters in an array are continuous filters. If tapped-delay lines are used to replace them, then referring to Fig. 4 the following expressions are obtained.

Let the i^{th} individual filter be

$$H_i(\omega) = \sum_{k=0}^{M} C_{ik}\, e^{-j\omega\Delta_{ik}} \tag{5.2-16}$$

where C_{ik} and Δ_{ik} are the weight and spacing at the k^{th} tap on the i^{th} filter.

The weights assume different values depending on the training environment. The post-summation filter $G(\omega)$ is fixed at all times by (2.2-9), i.e.,

$$G(\omega) = \phi_d^{-\frac{1}{2}}(\omega) \tag{5.2-17}$$

Substituting Eq. (5.2-16) and (5.2-17) into Eqs. (5.2-9) through (5.2-14), we obtain the following statistics

$$<y>_{S+N} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G|^2\, \underline{H}^T\, \underline{\Phi}_{xx}\, \underline{H}^*\, d\omega$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{k=0}^{M}\sum_{\ell=0}^{M} C_{ik}\, C_{h\ell}\, \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, \phi_d^{-1}\, \phi_{x_i x_h}\, e^{j\omega(\Delta_{h\ell} - \Delta_{ik})} \tag{5.2-18}$$

$$<y>_{N} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G|^2\, \underline{H}^T\, \underline{\Phi}_{nn}\, \underline{H}^*\, d\omega$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{k=0}^{M}\sum_{\ell=0}^{M} C_{ik}\, C_{k\ell}\, \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, \phi_d^{-1}\, \phi_{n_i n_h}\, e^{j\omega(\Delta_{h\ell} - \Delta_{ik})} \tag{5.2-19}$$

$$\bar{y}_{d.c.} = \langle y\rangle_{S+N} - \langle y\rangle_N = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G|^2 \underline{H}^T \underline{\Phi}_{ss} \underline{H}^* d\omega$$

$$= \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{k=0}^{M} \sum_{\ell=0}^{M} C_{ik} C_{h\ell} \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{j\omega(\tau_i-\tau_h)} e^{j\omega(\Delta_{h\ell}-\Delta_{ik})} \qquad (5.2\text{-}20)^{1}$$

$$\sigma_y^2 \cong \frac{1}{\pi T_{av}}\int_{-\infty}^{\infty} |G|^4 \{\underline{H}^T \underline{\Phi}_{nn} \underline{H}^*\}^2 d\omega$$

$$= \sum_{i=1}^{K} \sum_{i'=1}^{K} \sum_{h=1}^{K} \sum_{h'=1}^{K} \sum_{k=0}^{M} \sum_{k'=0}^{M} \sum_{\ell=0}^{M} \sum_{\ell'=0}^{M} C_{ik} C_{i'k'} C_{h\ell} C_{h'\ell'}$$

$$\frac{1}{\pi T_{av}}\int_{-\infty}^{\infty} d\omega\, \phi_d^{-2} \phi_{n_in_h} \phi_{n_{i'}n_{h'}} e^{j\omega(\Delta_{h\ell}+\Delta_{h'\ell'}-\Delta_{ik}-\Delta_{i'k'})} \qquad (5.2\text{-}21)$$

where $\phi_{x_ix_h}$ (or $\phi_{n_in_h}$) is the ij^{th} element of the input (or noise) spectral matrix and Eq. (2.2-21) is valid for the case of small signal-to-noise ratios.

It is readily seen that if every gain C_{ik} ($i = 1,2,\ldots k$ and $k = 0,1,\ldots M$) is multiplied by a constant such as the case of uncertain signal power, the final value of the output signal-to-noise ratio, defined as the change of dc level due to the appearance of a target signal divided by the rms fluctuation of the output, remains essentially unchanged.

5.3 Initial Behavior

Assuming the worst case where absolutely no information about the noise field is known, the gains associated with the input delayed by τ_i ($i = 1,2,\ldots K$) are set to 1 and the rest to zero so that a square-law detector is used at the starting moments. Here the output of each hydrophone is delayed to provide maxi-

---

[1] The integral appearing in Eq. (5.2-20) will, in general, yield delta-functions with infinite strength at certain instants. This difficulty does not arise in practice since most processes are bandlimited and the range of integration is ω_1 to ω_2, where ω_1 and ω_2 are finite numbers.

mum response in the signal direction, i.e.,

$$\underline{H}_1(\omega) = \underline{a}^* \tag{5.3-1}$$

The weights and spacings are simply

$$\begin{aligned} C_{ik} &= \delta_{ik} \\ \Delta_{ik} &= \tau_i \, \delta_{ik} \end{aligned} \tag{5.3-2}$$

Substituting Eq. (5.3-2) into Eq. (5.2-20) gives the dc change of the output due to the presence of target

$$\begin{aligned} \bar{y}^{(1)}_{d.c.} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega (\underline{H}_1^T \, \underline{\Phi}_{ss} \, \underline{H}_1^*) |G|^2 \\ &= \sum_{i=1}^{K} \sum_{h=1}^{K} C_{ii} \, C_{hk} \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \, e^{j\omega(\tau_i - \tau_h)} \, e^{j\omega(\tau_h - \tau_i)} \\ &= \sum_{i=1}^{K} \sum_{h=1}^{K} \frac{1}{2\pi} \int_{o}^{\omega_o} d\omega = \frac{K^2 \omega_o}{2\pi} \end{aligned} \tag{5.3-3}$$

The output variance is obtained by combining Eqs. (5.3-1), (5.3-2) and (5.2-21)

$$\begin{aligned} (\sigma_y^{(1)})^2 &= \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} d\omega \, (\underline{H}_1^T \, \underline{\Phi}_{nn} \, \underline{H}_1^*)^2 |G|^4 \\ &= \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{i'=1}^{K} \sum_{h'=1}^{K} \int_{o}^{\omega_o} d\omega \quad \phi_d^{-2} \, \phi_{n_i n_h} \, \phi_{n_{i'} n_{h'}} \, e^{j\omega(\tau_h + \tau_{h'} - \tau_i - \tau_{i'})} \\ &= \frac{1}{\pi T_{av}} \int_{o}^{\omega_o} \frac{N^2}{S^2} (K + \frac{I}{N} |\underline{a}^{*T} \, \underline{b}|^2)^2 d\omega \end{aligned} \tag{5.3-4}$$

a) The Output Signal-to-Noise Ratio

Dividing Eq. (5.3-3) by the square root of Eq. (5.3-4) gives the output signal-to-noise ratio

$$\mathrm{SNR}_1 = \frac{\frac{1}{2}\sqrt{\frac{T_{av}}{\pi}}\; K\,\omega_o\left(\frac{S}{N}\right)}{\left[\int_o^{\omega_o}\left(1 + \frac{I}{N}\,\frac{|\underline{a}^{*T}\,\underline{b}|^2}{K}\right)^2 d\omega\right]^{\frac{1}{2}}} \tag{5.3-5}$$

The term $\frac{|\underline{a}^{*T}\,\underline{b}|^2}{K}$ appearing in the numerator of Eq. (5.3-5) is proportional to the side lobe level in the direction of interference. For narrow band systems which have pronounced side lobe structure, the signal to-noise ratio is seen to depend on the side lobe level in the interference direction. This certainly agrees with our expectation.

We shall now evaluate the integral for the case of similar input spectra. Note that

$$\sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(\rho_h - \rho_i)} = \sum_{i=1}^{K}\sum_{h=1}^{K} \cos\omega\,(\rho_h - \rho_i)$$

$$= K + 2\sum_{i=1}^{K-1}\;\sum_{h=i+1}^{K} \cos\omega\,(\rho_h - \rho_i) \tag{5.3-6}$$

The value of the double sum in Eq. (5.3-6) can be further evaluated for our case of a linear array with equal spaced hydrophones. If such an array is steered broadside; i.e., if the target is at a location perpendicular to the array axis, then

$$\rho_h - \rho_i = |h - i|\,\frac{d}{c}\sin\theta_I = |h - i|\,\rho_o \tag{5.3-7}$$

and the double sum in Eq. (5.3-6) can be replaced by a single sum

$$\sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(\rho_h - \rho_i)} = K + 2\sum_{i=1}^{K-1}(K - i)\cos\omega\, i\,\rho_o \tag{5.3-8}$$

Using Eq. (5.3-8), we have

$$(\sigma_y^{(1)})^2 = (\frac{N}{S})^2 \frac{K}{\pi T_{av}} \int_0^{\omega_o} (1 + \frac{I}{N} \frac{|\underline{a}^{*T}\underline{b}|^2}{K})^2 \, d\omega$$

$$= (\frac{N}{S})^2 \frac{K}{\pi T_{av}} \int_0^{\omega_o} d\omega \, [1 + \frac{I}{KN} \sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(\rho_h - \rho_i)}]^2$$

$$= (\frac{N}{S})^2 \frac{K}{\pi T_{av}} \int_0^{\omega_o} \{1 + \frac{I}{N} [K + 2 \sum_{i=1}^{K-1} (K-i) \cos \omega i \rho_o]\}^2 \, d\omega$$

$$= (\frac{N}{S})^2 \frac{K\omega_o}{\pi T_{av}} \Big\{ 1 + \frac{I}{N} [2 + \frac{4}{K} \sum_{i=1}^{K-1} \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o} (K-i)$$

$$+ \frac{I^2}{N^2} \{1 + \frac{4}{K} \sum_{i=1}^{K-1} \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o} (K-i)$$

$$+ \frac{2}{K^2} \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} [\frac{\sin \omega_o (i-h)\rho_o}{\omega_o (i-h)\rho_o} + \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o}](K-i)(K-h)\} \Big\}$$

(5.3-9)

In most practical cases the maximum frequency processed is very high such that $\omega_o \rho_o >> 1$. Then the sums associated with ρ_o make negligible contribution except for $i = h$ and we have a simpler expression

$$\int_0^{\omega_o} (1 + \frac{I}{N} \frac{|\underline{a}^{*T}\underline{b}|^2}{K})^2 \, d\omega$$

$$\approx \omega_o \{1 + 2 \frac{I}{N} + \frac{I^2}{N^2} [1 + \frac{2}{K^2} \sum_{i=1}^{K-1} (K-i)^2]\}$$

$$= \omega_o \{1 + 2 \frac{I}{N} + \frac{I^2}{N^2} (\frac{2}{3} K + \frac{1}{3K})\} \tag{5.3-10}$$

The output signal-to-noise ratio becomes

$$SNR_1 \approx \frac{1}{2} (\frac{T_{av} \omega_o}{\pi})^{\frac{1}{2}} \frac{S}{N} K \{1 + 2 \frac{I}{N} + \frac{I^2}{N^2} (\frac{2}{3}K + \frac{1}{3K})\}^{-\frac{1}{2}} \tag{5.3-11}$$

For most cases of practical interest, the number of hydrophones in an array is large $K >> 1$ so that for ambient-noise-dominated environment

$$SNR_1 \; \alpha \; K \left(\frac{S}{N}\right) \quad \text{when} \quad \left(\frac{N^2}{I}\right) >> \frac{2}{3} K \tag{5.3-12}$$

and for interference-dominated environment

$$SNR_1 \; \alpha \; K^{\frac{1}{2}} \left(\frac{S}{I}\right) \quad \text{when} \quad \left(\frac{N}{I}\right)^2 << \frac{2}{3} K \tag{5.3-13}$$

The results concerning the output signal-to-noise ratio have been previously derived by Schulthiess [32] for a conventional power detector under the assumptions that the interference and ambient-noise are white over $0 < \omega < \omega_o$ and $G(\omega) = 1$. In our case $G(\omega) = \phi_d^{-\frac{1}{2}}$ and the input spectra are similar rather than constant over the same frequency range. This is equivalent to inserting an Eckart filter

$$|G_E(\omega)|^2 = \frac{\phi_d(\omega)}{\phi_n^2(\omega)} \tag{5.3-14}$$

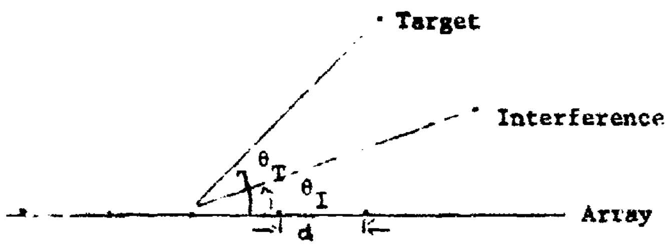
after the beamforming point in the absence of interference. The effect of inserting this filter has also been considered by Schultheiss and reported elsewhere [39].

b) <u>Directivity Patterns</u>

The average output of the squarer, $\overline{y}$, yields the so-called directivity pattern which may be obtained by varying the electrical time delays and keeping the physical orientation of the array fixed, or by keeping the electrical time delays fixed and varying the physical orientation of the array relative to the plane wave signal. It is a function of the target bearing relative to the bearing angle of the major lobe of the array pattern.

Let θ_T and θ_I be the target and interference bearings relative to the broadside condition with the convention of signs that angles are measured clock-

wise from broadside.



If we define the following terms for a linear array of equally spaced hydrophones,

$$\tau_o = \frac{d}{c} \sin \theta_T \tag{5.3-15}$$

$$\rho_o = \frac{d}{c} \sin \theta_I \tag{5.3-16}$$

$$\hat{\tau}_o = \frac{d}{c} \sin \theta \tag{5.3-17}$$

d = hydrophone spacing

c = velocity of sound in water

then the signal and interference delays at the i^{th} phone in a linear array with equal spaced hydrophones would be

$$\tau_i = (K-i)\tau_o \tag{5.3-18}$$

$$\rho_i = (K-i)\rho_o \tag{5.3-19}$$

and the steering angle

$$\hat{\tau}_i = (K-i)\hat{\tau}_o \tag{5.3-20}$$

where $\hat{\tau}_o$, the looking angle appearing in the individual filters, is the independent variable of the directivity pattern.

It is seen from Eq. (5.2-18) that the averaged output is

$$\bar{y}_1 = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \; |G|^2 \; (\hat{\underline{H}}_1 \; \underline{\Phi}_{xx} \; \hat{\underline{H}}^*)$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K} \frac{1}{2\pi}\int_{0}^{\omega_o} d\omega \; \phi_d^{-1} \; [\phi_d \, e^{j\omega(\tau_i - \tau_h)} + \phi_n \, \delta_{ih}$$

$$+ \phi_I \, e^{j\omega(\rho_i - \rho_h)}] \; e^{j\omega(\hat{\tau}_h - \hat{\tau}_i)}$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K} \frac{1}{2\pi}\int_{0}^{\omega_o} d\omega [e^{j\omega(i-h)(\tau_o - \hat{\tau}_o)}$$

$$+ \frac{I}{S} \, e^{j\omega(i-h)(\rho_o - \hat{\tau}_o)}]$$

$$+ K \frac{N}{S} \frac{1}{2\pi}\int_{0}^{\omega_o} d\omega \qquad (5.3\text{-}21)$$

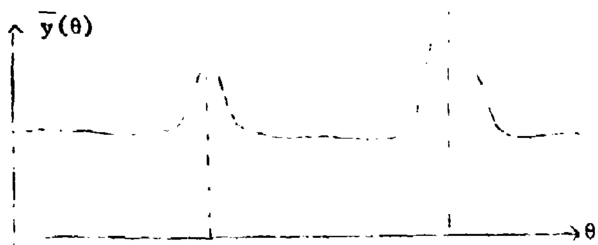
Upon using Eqs. (5.3-6) and (5.3-8) and carrying out the integration, we have

$$\bar{y}_1(\theta) = \frac{\omega_o}{2\pi} \quad K(1 + \frac{N}{S} + \frac{I}{S})$$

$$+ 2 \sum_{i=1}^{K-1}(K-i) \; \frac{\sin \omega_o i(\tau_o - \hat{\tau}_o)}{\omega_o i(\tau_o - \hat{\tau}_o)}$$

$$+ \frac{I}{S} \; \frac{\sin \omega_o i(\rho_o - \hat{\tau}_o)}{\omega_o i(\rho_o - \hat{\tau}_o)} \qquad (5.3\text{-}22)$$

In the above $\rho_o = \frac{d}{c}\sin\theta_I$, $\tau_o = \frac{d}{c}\sin\theta_T$, $\hat{\tau}_o = \frac{d}{c}\sin\theta$, θ being the independent variable in calculating the directivity pattern $\bar{y}(\theta)$. If target and interference are well separated in bearing, the directivity pattern will take the general form shown below because of the plus signs appearing in Eq. (5.3-22).



In the signal direction $\hat{\tau} = \tau_o$ and for $\omega_o \rho_o >> 1$

$$\bar{y}_1(\theta = \theta_T) \simeq \frac{\omega_o}{2\pi} K \left[\left(\frac{N}{S} + \frac{I}{S}\right) + K \right] \tag{5.3-23}$$

Similarly in the interference direction $\hat{\tau}_o = \rho_o$

$$\bar{y}_1(\theta = \theta_I) \simeq \frac{\omega_o}{2\pi} K \left[\left(1 + \frac{N}{S}\right) + K \frac{I}{S} \right] \tag{5.3-24}$$

and in any other directions

$$\bar{y}_1(\theta) \simeq \frac{\omega_o}{2\pi} K\left(1 + \frac{N}{S} + \frac{I}{S}\right) \tag{5.3-25}$$

5.4 Final Behavior

a) Optimum Gains and Training Environment

The final form of our adaptive array processor is the one in which all the gains are set at their optimum values. From the convergence properties of adaptive tapped-delay-line filters we know that the final values of the gains are different under different training environment. They are

$$\underline{W}_\infty^{(S+N)} = (\underline{R}_\nu + \underline{R}_\xi)^{-1} \underline{R}_{d\xi} \tag{5.4-1}$$

$$\underline{W}_\infty^{(N)} = \underline{R}_\nu^{-1} \underline{R}_{d\xi} \tag{5.4-2}$$

where

$$\underline{W}^T = [c_{10}\ c_{11} \cdots c_{1M}\ c_{20} \cdots c_{2M} \cdots c_{K0} \cdots c_{KM}]$$

$$\underline{R}_\xi = E[\underline{\xi}\ \underline{\xi}^T]$$

$$\underline{R}_\nu = E[\underline{\nu}\ \underline{\nu}^T]$$

$$\underline{\xi}^T = [s_1(t)\ s_1(t - \Delta) \cdots s_K(t - M\Delta)]$$

$$\underline{\nu}^T = [n_1(t)\ n_1(t - \Delta) \cdots n_K(t - M\Delta)]$$

as indicated in Fig. 4.

It has been shown in Sect. 2.4 that for the tapped-delay-line filters, if

the gains are set according to Eq. (5.4-1) or (5.4-2), they are approximately equivalent to the optimum filters

$$\underline{H}_{\infty}^{(S+N)} = (\underline{\Phi}_{ss}^{*} + \underline{\Phi}_{nn}^{*})^{-1} \phi_d \, \underline{a}^{*} \tag{5.4-3}$$

$$\underline{H}_{\infty}^{(N)} = \underline{\Phi}_{nn}^{*-1} \phi_d \, \underline{a}^{*} \tag{5.4-4}$$

The accuracy of the above approximations depend on the number of taps used. For the case of similar input spectra we shall see that filters defined by Eq.(5.4-3) or (5.4-4) can be realized completely by tapped delay lines with proper settings and proper spacings.

Although the optimum gains are difficult to be expressed analytically using Eq. (5.4-1) or (5.4-2), they can be computed in the frequency domain by

$$C_{ik} = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_i(\omega) \, e^{j\omega t} \, d\omega \Bigg|_{t = k\Delta} \tag{5.4-5}$$

for the k^{th} gain on the i^{th} filter. $H_i(\omega)$ is just the i^{th} row of either Eq. (5.4-3) or (5.4-4).

We shall first of all consider Eq. (5.4-4).

Since

$$\underline{H}_{\infty}^{(N)} = \underline{\Phi}_{nn}^{*-1} \phi_d \, \underline{a}^{*} = \frac{\phi_d}{\phi_n} [\underline{I} - \frac{\underline{b}^{*} \underline{b}^{T}}{K + \phi_n/\phi_I}] \, \underline{a}^{*} \tag{5.4-6}$$

its i^{th} row is

$$H_i^{(N)} = \frac{\phi_d}{\phi_n} [a_i^{*} - \frac{b_i^{*} \sum_{k=1}^{K} b_k \, a_k^{*}}{K + \phi_n/\phi_I}] \tag{5.4-7}$$

so that the impulse response is

$$h_i(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi_d}{\phi_n} [e^{-j\omega\tau_i} - \frac{e^{-j\omega\rho_i} \sum_{k=1}^{K} e^{j\omega(\rho_k - \tau_k)}}{K + \phi_n/\phi_I}] \, e^{j\omega t} \, dt$$

$$= \frac{S}{N}\ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega(t-\tau_i)} d\omega$$

$$- \frac{S}{N}\ \frac{1}{K + N/_I}\ \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \sum_{k=1}^{K} e^{j\omega(t - \rho_i + \rho_k - \tau_k)} \tag{5.4-8}$$

as is shown in Section 2.5, Eqs. (2.5-12), (2.5-13)

$$C_{ik}^{(N)} = \frac{S}{N} (\delta_{ik} - \frac{1}{K + \frac{N}{I}}) \tag{5.4-9}$$

at

$$\Delta_{ik} = \rho_i - \rho_k + \tau_k \tag{5.4-10}$$

If the gains are adjusted in the presence of target, we shall use Eq.(5.4-3) instead of Eq. (5.4-4).

Since

$$(\underline{\Phi}_{ss}^{*} + \underline{\Phi}_{nn}^{*})^{-1}\ \phi_d\ \underline{a}^{*} = (\underline{\Phi}_{nn}^{*} + \phi_d\ \underline{a}^{*}\ \underline{a}^{T})^{-1}\ \phi_d\ \underline{a}^{*}$$

$$= [\underline{\Phi}_{nn}^{*-1} - \frac{\phi_d\ \underline{\Phi}_{nn}^{*-1}\ \underline{a}^{*}\ \underline{a}^{T}\ \underline{\Phi}_{nn}^{*-1}}{1 + \phi_d\ \underline{a}^{T}\ \underline{\Phi}_{nn}^{*-1}\ \underline{a}^{*}}]\ \phi_d\ \underline{a}^{*} = \frac{\underline{\Phi}_{nn}^{*-1}\ \phi_d\ \underline{a}^{*}}{1 + \phi_d\ \underline{a}^{T}\ \underline{\Phi}_{nn}^{*-1}\ \underline{a}^{*}} \tag{5.4-11}$$

and let

$$1 + \phi_d\ \underline{a}^{T}\ \underline{\Phi}_{nn}^{*-1}\ \underline{a}^{*} = 1 + \frac{\phi_d}{\phi_n} [K - \frac{|\underline{a}^{*T}\underline{b}|^2}{K + \phi_{n}/\phi_I}]$$

$$\triangleq \frac{1}{K_1} \tag{5.4-12}$$

we have

$$\underline{H}_{\infty}^{(S+N)} = \underline{H}_{\infty}^{(N)}\ K_1 \tag{5.4-13}$$

For the case of similar input spectra, K_1 is just a constant

$$K_1 = \frac{N^2 + KNI}{N^2 + KNI + KSN + SI(K^2 - K_2^2)} \tag{5.4-14}$$

where

$$K_2^2 \triangleq |\underline{a}^{*T}\underline{b}| \leq K^2 \tag{5.4-15}$$

with equality when $\underline{a} = \underline{b}$ or when target and interference are in the same direction.

Therefore

$$C_{ik}^{(S+N)} = K_1 C_{ik}^{(N)} \quad \text{for } i = 1,2,\cdots, K;\ k = 1,2,\cdots,K \tag{5.4-16}$$

In other words, for similar input spectra the optimal gains trained under noise alone differ from those under noise plus signal only by a multiplicative constant K_1. Thus, all the performance criteria (output signal-to-noise ratio, and directivity pattern) remain essentially unchanged regardless of the training environment. Of course, should the input spectra have distinctive shapes, $C_{ik}^{(S+N)}$ and $C_{ik}^{(N)}$ would assume different values. If the signal-to-noise ratio is small at the input to the squarer

$$\phi_d\, \underline{a}^{*T}\, \underline{\phi}_{nn}^{*-1}\, \underline{a} << 1$$

the constant K_1 is close to unity independent of the input spectral shapes.

In view of the above discussions we shall use the optimal gains defined by Eq. (5.4-9) in analyzing the final behavior of our adaptive processor.

b) The Output Signal-to-noise Ratio

From Eqs. (5.4-9), (5.4-10) and (5.2-20), the dc change of the output is

$$\overline{y}_{d.c}^{\infty} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G|^2\, (\underline{H}_{\infty}^T\, \underline{\Phi}_{ss}\, \underline{H}_{\infty}^*)\,d\omega$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{k=1}^{K}\sum_{\ell=1}^{K} \frac{S}{N}\left(\delta_{ik} - \frac{1}{K+N/_I}\right)\frac{S}{N}\left(\delta_{h\ell} - \frac{1}{K+N/_I}\right)$$

$$\cdot \frac{1}{2\pi}\int_{0}^{\omega_o} d\omega\, e^{j\omega(\rho_h - \rho_\ell + \tau_\ell - \rho_i + \rho_k - \tau_k)} \tag{5.4-17}$$

For the sake of simplicity, we shall assume that the array is steered broadside. Then $\tau_i = 0$ $(i = 1,2,\cdots,K)$ and the above expression reduces to

$$\overline{y}^{(\infty)}_{d.c} = (\frac{S}{N})^2 \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{k=1}^{K} \sum_{\ell=1}^{K} (\delta_{ik} - \frac{1}{K+N/_I})(\delta_{h\ell} - \frac{1}{K+N/_I})$$

$$\cdot \frac{1}{2\pi} \int_0^{\omega_o} d\omega \; e^{j\omega(\rho_h - \rho_\ell - \rho_i + \rho_k)}$$

$$= (\frac{S}{N})^2 \frac{\omega_o}{2\pi} \Big\{ K^2 - \frac{2K}{K+N/_I} [K + 2 \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o}]$$

$$+ \frac{1}{(K+N/I)^2} [K^2 + 4K \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o}]$$

$$+ 2 \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} (K-i)(K-h) (\frac{\sin \omega_o (i-h)\rho_o}{\omega_o (i-h)\rho_o} + \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o})] \Big\} \tag{5.4-18}$$

or

$$\overline{y}^{\infty}_{d.c} = \frac{\omega_o}{2\pi} (\frac{S}{N})^2 \frac{K^2 (K-1 + N/_I)^2}{(K + N/_I)^2}$$

$$\cdot \Big\{ 1 - \frac{4}{K(K-1 + \frac{N}{I})} \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o}$$

$$+ \frac{2}{K^2 (K-1 + N/_I)^2} \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} (K-i)(K-h) [\frac{\sin \omega_o (i-h)\rho_o}{\omega_o (i-h)\rho_o}$$

$$+ \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o}] \Big\} \tag{5.4-19}$$

Let us now consider the variance of the detector output. From Eqs. (5.4-9), (5.4-10) and (5.2-2) we have

$$(\sigma_y^{(\infty)})^2 = \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4 \{\underline{H}_\infty^T \Phi \underline{H}^*\}^2 d\omega$$

$$= \sum_{i=1}^{K} \sum_{i'=1}^{K} \sum_{h=1}^{K} \sum_{h'=1}^{K} \sum_{k=1}^{K} \sum_{k'=1}^{K} \sum_{\ell=1}^{K} \sum_{\ell'=1}^{K}$$

$$\cdot \left(\frac{S}{N}\right)^4 \left(\delta_{ik} - \frac{1}{K+N/_I}\right)\left(\delta_{i'k'} - \frac{1}{K+N/_I}\right)\left(\delta_{h\ell} - \frac{1}{K+N/_I}\right)\left(\delta_{h'\ell'} - \frac{1}{K+N/_I}\right)$$

$$\cdot \frac{1}{\pi T_{av}} \int_o^{\omega_o} d\omega \; \phi_d^{-2} \; \left(\phi_n \, \delta_{ih} + \phi_I \, e^{j\omega(\rho_i - \rho_h)}\right)$$

$$\cdot \left(\phi_n \, \delta_{i'h'} + \phi_I \, e^{j\omega(\rho_{i'} - \rho_{h'})}\right)$$

$$\cdot \, e^{j\omega(\rho_h - \rho_\ell + \tau_\ell + \rho_{h'} - \rho_{\ell'} + \tau_{\ell'} - \rho_i + \rho_k - \tau_k - \rho_{i'} + \rho_{k'} - \tau_{k'})}$$

(5.4-20)

Rearranging,

$$\left(\sigma_y^{(\infty)}\right)^2 = \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{k=1}^{K} \sum_{\ell=1}^{K} \frac{1}{\pi T_{av}} \int_o^{\omega_o} d\omega \; \phi_d^{-2} \left\{ \frac{S}{N} \left(\delta_{ik} - \frac{1}{K+N/_I}\right)\right.$$

$$\cdot \left[\phi_n \, \delta_{ih} + \phi_I \, e^{j\omega(\rho_i - \rho_h)}\right] \frac{S}{N} \left(\delta_{h\ell} - \frac{1}{K+N/_I}\right)$$

$$\left. \cdot \, e^{j\omega(\rho_h - \rho_\ell + \tau_\ell - \rho_i + \rho_k - \tau_k)} \right\}^2$$

$$= \frac{1}{\pi T_{av}} \int_o^{\omega_o} d\omega \left\{ \sum_{i=1}^{K} \sum_{h=1}^{K} \left[\frac{S}{N} \left(e^{-j\omega\tau_i} - \frac{1}{K+N/_I} \sum_{k=1}^{K} e^{j\omega(-\rho_i + \rho_k - \tau_k)}\right]\right.\right.$$

$$\cdot \left[\frac{S}{N} \left(e^{-j\omega\tau_h} - \frac{1}{K+N/_I} \sum_{\ell=1}^{K} e^{j\omega(\rho_h - \rho_\ell + \tau_\ell)}\right]\right.$$

$$\left. \cdot \left[\frac{N}{S} \, \delta_{ih} + \frac{I}{S} \; e^{j\omega(\rho_i - \rho_h)}\right] \right\}^2$$

$$= \frac{1}{\pi T_{av}} \int_o^{\omega_o} d\omega \; \left(\frac{S}{N}\right)^2 \left\{ K - \frac{1}{K+N/_I} \left| \sum_{i=1}^{K} e^{j\omega(\rho_i - \tau_i)} \right|^2 \right\}^2$$

(5.4-21)

For $\tau_i = 0$, it reduces to

$$(\sigma_y^{(\infty)})^2 = \frac{1}{\pi T_{av}} (\frac{S}{N})^2 \int_o^{\omega_o} d\omega \left\{ K^2 - \frac{2K}{K+N/_I} \sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(\rho_i - \rho_h)} + \frac{1}{(K+N/_I)^2} [\sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(\rho_i - \rho_h)}]^2 \right\}$$

$$= \frac{\omega_o}{\pi T_{av}} (\frac{S}{N})^2 \left\{ K^2 - \frac{2K}{K+N/_I} [K + 2 \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i\rho_o}{\omega_o i \rho_o}] + \frac{1}{(K+N/_I)^2} [K^2 + 4K \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o} + 2 \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} (K-i)(K-h) (\frac{\sin \omega_o (i-h)\rho_o}{\omega_o (i-h)\rho_o} + \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o})] \right\} \qquad (5.4\text{-}22)$$

Dividing Eq. (5.4-19) by the square root of Eq. (5.4-22) yields the final form of the output signal-to-noise ratio

$$SNR_\infty = \frac{1}{2} (\frac{T_{av}\, \omega_o}{\pi})^{\frac{1}{2}} \frac{S}{N} \frac{K(K-1 + N/_I)}{K + N/_I} \cdot \left\{ 1 - \frac{4}{K(K-1 + N/_I)} \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o i \rho_o}{\omega_o i \rho_o} + \frac{2}{K^2 (K-1+N/_I)^2} \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} (K-i)(K-h) \cdot [\frac{\sin \omega_o (i-h)\rho_o}{\omega_o (i-h)\rho_o} + \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o}] \right\}^{\frac{1}{2}} \qquad (5.4\text{-}23)$$

If $\omega_o \rho_o >> 1$ then

$$SNR_\infty \simeq \frac{1}{2} (\frac{T_{av}\, \omega_o}{\pi})^{\frac{1}{2}} \frac{S}{N} (K-1) \frac{1+N/(K-1)I}{1+N/KI} [1 + \frac{1}{3} \frac{(K-1)(2K-1)}{K(K-1+N/_I)^2}]^{\frac{1}{2}}$$

$$\geq \frac{1}{2} (\frac{T_{av}\, \omega_o}{\pi})^{\frac{1}{2}} \frac{S}{N} (K-1) \qquad (5.4\text{-}24)$$

Eq. (5.4-23) or (5.4-24) gives the asymptotic performance of the adaptive array processor. Since the training period is definitely of finite time, the actual signal-to-noise ratio is lower than that given by Eq. (5.4-23) or (5.4-24). It is to be noted that Eq. (5.4-23) is just the output signal-to-noise ratio of an optimal (likelihood ratio) detector first investigated by Schultheiss [32] and then by Tuteur [33] from a simpler formulation under the same assumption [similar input spectra over $(0, \omega_o)$]. This is no coincidence but stems from the fact that adequately large number of taps is used (number of taps per hydrophone output = number of hydrophones in an array) and that for similar input spectra the continuous individual filters are just combinations of time delays and constants. Therefore, it is reasonable to expect that in this special case the adaptive array processor can, in principle, converge to the optimal processor after sufficiently long period of adaptation.

In the absence of interference the behavior of the processor remains unaltered throughout the training period because

$$SNR_1 = SNR_\infty = \frac{1}{2}\left(\frac{T_{av}\,\omega_o}{\pi}\right)^{\frac{1}{2}} \frac{S}{N} K \text{ for } I = 0 \tag{5.4-25}$$

Without interference we do not take advantage of the reason why the optimum detector is superior to the conventional detector:

1) It can combat noise correlation between hydrophones and

2) It can utilize variations in input signal-to-noise ratio over the processed frequency band.

Since from Eqs. (5.4-19) and (5.4-21) we can write

$$SNR_\infty = \frac{1}{2}\sqrt{\frac{T_{av}}{\pi}} \left[\int_0^{\omega_o} \left(\frac{S}{N}\right)^2 \left[K - \frac{|\underline{a}^{*T}\underline{b}|^2}{K+N/_I}\right]^2 d\omega\right]^{\frac{1}{2}} \tag{5.4-26}$$

we see in the above equation that in the absence of the interference or for very large values of N/I the side lobe factor $\frac{|\underline{a}^{*T}\underline{b}|^2}{K+N/_I}$ approaches zero and output signal-to-noise ratio depends linearly on the number of hydrophones K. For very

strong interference ($N/I << 1$), the same quantity approaches to $|\underline{a}^{*T}\underline{b}|^2/K$. The size of this side lobe in a fixed direction depends on frequency. In a narrow band system large variations in this term may occur. That is, there will be distinct maxima and nulls in the side lobe pattern. In a broad band system the magnitude of the side lobe is averaged over frequency.

c) Directivity Pattern

Using Eq. (5.2-18), (5.4-9) and (5.4-10), we have

$$\bar{y}_\infty = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G|^2\, \hat{\underline{H}}_\infty^T\, \underline{\Phi}_{xx}\, \hat{\underline{H}}^*\, d\omega$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\left(\frac{S}{N}\right)^2 \left(\delta_{ik} - \frac{1}{K+N/_I}\right)\left(\delta_{h\ell} - \frac{1}{K+N/_I}\right)$$

$$\cdot \frac{1}{2\pi}\int_0^{\omega_o} d\omega\, \phi_d^{-1}\, [\phi_d\, e^{j\omega(\tau_i - \tau_h)} + \phi_n\, \delta_{ih}$$

$$+ \phi_I\, e^{j\omega(\rho_i - \rho_h)}]\, e^{j\omega(\rho_h - \rho_\ell + \hat{\tau}_\ell - \rho_i + \rho_k - \hat{\tau}_k)}$$

$$= \frac{1}{2\pi}\int_0^{\omega_o}\left\{\sum_{i=1}^{K}\frac{S}{N}\, [e^{-j\omega\hat{\tau}_i} - \frac{e^{-j\omega\rho_i}}{K+N/_I}\sum_{k=1}^{K} e^{j\omega(\rho_k - \hat{\tau}_k)}]\, e^{j\omega\tau_i}\right\}^2 d\omega$$

$$+ \frac{1}{2\pi}\int_0^{\omega_o} d\omega\left(\frac{S}{N}\, e^{-j\omega\hat{\tau}_i} - \frac{e^{-j\omega\rho i}}{K+N/I}\sum_{k=1}^{K} e^{j\omega(\rho_k - \hat{\tau}_k)}\right] e^{j\omega\hat{\tau}_i} \qquad (5.4\text{-}27)$$

The writing of the above expression is permissible because for

$\underline{H}_\infty = \underline{\Phi}_{nn}^{*-1}\, \phi_d\, \underline{a}^*$ we have

$$|G|^2\, \hat{\underline{H}}_\infty^T\, \underline{\Phi}_{xx}\, \hat{\underline{H}}^* = \phi_d^{-1}\, \hat{\underline{a}}^{*T}\, \underline{\Phi}_{nn}^{-1}\, \phi_d\, (\phi_d\, \underline{a}\, \underline{a}^{*T} + \underline{\Phi}_{nn})\underline{\Phi}_{nn}^{-1}\, \hat{\underline{a}}\, \phi_d$$

$$= \phi_d^2\, (\hat{\underline{a}}^{*T}\, \underline{\Phi}_{nn}^{-1}\, \underline{a})^2 + \phi_d\, \hat{\underline{a}}^{*T}\, \underline{\Phi}_{nn}^{-1}\, \hat{\underline{a}} \qquad (5.4\text{-}28)$$

which is equivalent to the integrand appearing in Eq. (5.4-27).

Eq. (5.4-27) is further simplified to

$$\overline{y}_\infty = \frac{1}{2\pi}\int_0^{\omega_o} d\omega \left(\frac{S}{N}\right)^2 \left\{ \sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(h-i)(\tau_o - \hat{\tau}_o)} \right.$$

$$- \frac{2}{K+N/I} \sum_{i=1}^{K} e^{j\omega(\tau_i - \hat{\tau}_i)} \sum_{h=1}^{K}\sum_{k=1}^{K} e^{j\omega(\hat{\tau}_h + \rho_h - \rho_k + \tau_k)}$$

$$\left. + \frac{1}{(K+N/I)^2} \left[\sum_{i=1}^{K} e^{j\omega(\rho_i - \hat{\tau}_i)} \sum_{h=1}^{K} e^{j\omega(\tau_h - \rho_h)} \right]^2 \right\}$$

$$+ \frac{1}{2\pi}\int_0^{\omega_o} d\omega \left(\frac{S}{N}\right) \left[K - \frac{1}{K+N/I} \sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(i-h)(\hat{\tau}_o - \rho_o)}\right] \tag{5.4-29}$$

Since

$$\sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{k=1}^{K} e^{j\omega(\tau_i - \hat{\tau}_i - \hat{\tau}_h + \rho_h - \rho_k + \tau_k)}$$

$$= K \sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(i-h)(\tau_o - \hat{\tau}_o)}$$

$$+ \sum_{i=1}^{K}\sum_{h=1}^{K}\sum_{\substack{k=1 \\ h \neq k}}^{K} e^{j\omega(\tau_i - \hat{\tau}_i - \hat{\tau}_h + \rho_h - \rho_k + \tau_k)} \tag{5.4-30}$$

and

$$\left(\sum_{i=1}^{K} e^{j\omega(\rho_i - \hat{\tau}_i)} \sum_{h=1}^{K} e^{j\omega(\tau_h - \rho_h)} \right)^2$$

$$= \sum_{i=1}^{K}\sum_{h=1}^{K} e^{j\omega(i-h)(\rho_o - \hat{\tau}_o)} \sum_{k=1}^{K}\sum_{\ell=1}^{K} e^{j\omega(k-\ell)(\tau_o - \rho_o)}$$

$$= \left[K + 2 \sum_{i=1}^{K-1} (K-i) \cos \omega\, i(\rho_o - \hat{\tau}_o)\right]$$

$$\cdot \left[K + 2 \sum_{h=1}^{K-1} (K-h) \cos \omega\, h(\rho_o - \hat{\tau}_o)\right]$$

$$= K^2 + 2K \sum_{i=1}^{K-1} (K-i)\left[\cos \omega\, i(\rho_o - \hat{\tau}_o) + \cos \omega\, i(\tau_o - \rho_o)\right]$$

$$+ 2 \sum_{i=1}^{K-1} (K-i)^2 \cos \omega\, i(\tau_o - \hat{\tau}_o)$$

$$+ 2 \sum_{\substack{i=1 \\ i \neq h}}^{K-1} \sum_{h=1}^{K-1} (k-i)(K-h) \Big\{ \cos (i\rho_o - i\hat{\tau}_o + h\tau_o - h\rho_o)$$

$$+ \cos (i\rho_o - i\hat{\tau}_o - h\tau_o + h\rho_o)] \Big\} \qquad (5.4\text{-}31)$$

We have

$$\overline{y}_\infty \simeq \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \Bigg\{ K + 2 \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o\, i(\tau_o - \hat{\tau}_o)}{\omega_o\, i(\tau_o - \hat{\tau}_o)}$$

$$- \frac{2K}{K+N/I} \left[K + 2 \sum_{i=1}^{K-1} (K-1) \frac{\sin \omega_o\, i(\tau_o - \hat{\tau}_o)}{\omega_o\, i\,(\tau_o - \hat{\tau}_o)}\right]$$

$$+ \frac{1}{(K+N/I)^2} \Bigg[K^2 + 2K \sum_{i=1}^{K-1} (K-i) \Bigg(\frac{\sin \omega_o\, i(\rho_o - \hat{\tau}_o)}{\omega_o\, i(\rho_o - \hat{\tau}_o)}$$

$$+ \frac{\sin \omega_o\, i(\tau_o - \rho_o)}{\omega_o\, i(\tau_o - \rho_o)}\Bigg) + 2 \sum_{i=1}^{K-1} (K-i)^2 \frac{\sin \omega_o\, i(\tau_o - \hat{\tau}_o)}{\omega_o\, i(\tau_o - \hat{\tau}_o)}\Bigg] \Bigg\}$$

$$+ \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)\left[\frac{K(K-1+N/I)}{(K+N/I)} - 2 \frac{1}{K+N/I} \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o\, i(\hat{\tau}_o - \rho_\rho)}{\omega_o\, i(\hat{\tau}_o - \rho_o)}\right] \qquad (5.4\text{-}32)$$

In the above the last terms on the right-hand sides of (5.4-30) and (5.4-31) have been omitted because these terms always contribute very little upon integra-

tion. They are divided by ω_o which is large to make $\frac{\sin \omega_o \tau}{\omega_o \tau}$ negligible. At an angle far away from both the signal and interference directions, we may neglect all the oscillatory terms and get

$$\bar{y}_\infty(\theta) = \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \frac{K^2(K-1+N/I)^2}{(K+N/I)^2} - \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 K(K-1)$$

$$+ \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right) \frac{K(K-1+N/I)}{(K+N/I)} \qquad (5.4\text{-}33)$$

In the signal direction, $\hat{\tau}_o = \tau_o$

$$\bar{y}_\infty(\theta = \theta_T) \simeq \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \frac{K^2(K-1+N/I)^2}{(K + N/I)^2}$$

$$+ \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \frac{K^2(K-1+N/I)^2}{(K+N/I)^2} \frac{1}{3} \frac{(K-1)(2K-1)}{K(K-1+N/I)}$$

$$+ \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right) \frac{K(K-1+N/I}{(K+N/I)} \qquad (5.4\text{-}34)$$

and in the interference direction, $\hat{\tau}_o = \rho_o$

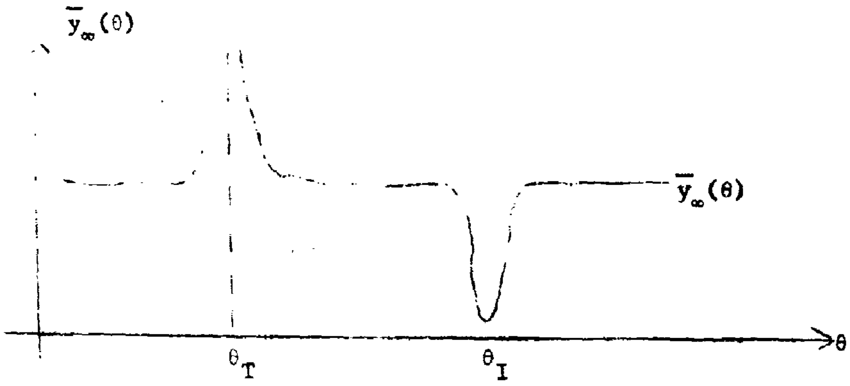
$$\bar{y}_\infty(\theta = \theta_I) \simeq \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \frac{K^2(K-1+N/I)^2}{(K+N/I)^2}$$

$$- \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 K(K-1) + \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right) \frac{K(K-1+N/I)}{(K+N/I)}$$

$$- \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right)^2 \frac{(K-1)}{(K+N/KI)^2} - \frac{\omega_o}{2\pi} \left(\frac{S}{N}\right) \frac{K(K-1)}{K+N/I} \qquad (5.4\text{-}35)$$

Although exact shapes of the optimal directivity pattern given by Eq. (5.4-32) cannot be plotted without assuming specific input power levels, they will take the general form shown below by comparing Eqs. (5.4-33) through

(5.4-35). **There is always a maximum in the signal direction and a minimum in the interference direction. Specific results will be given in the next chapter.**



5.5 Adaptive Behavior

The behavior of a tapped delay line filter is completely determined once the weights and spacings are known. If C_{ik} and Δ_{ik} are the weight and spacing at the k^{th} tap on the i^{th} filter, then for an array of K hydrophones and M taps at each sensor output, the impulse response vector of the pre-processor (from hydrophone outputs to the beamformer) is

$$\underline{h}(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \\ \cdot \\ \cdot \\ \cdot \\ h_K(t) \end{bmatrix} = \sum_{\ell=0}^{M} \begin{bmatrix} C_{1\ell}\ \delta(t - \Delta_{1\ell}) \\ C_{2\ell}\ \delta(t - \Delta_{2\ell}) \\ \cdot \\ \cdot \\ \cdot \\ C_{K\ell}\ \delta(t - \Delta_{K\ell}) \end{bmatrix} \tag{5.5-1}$$

and the transfer function vector

$$\underline{H}(\omega) = \begin{bmatrix} H_1(\omega) \\ H_2(\omega) \\ \cdot \\ \cdot \\ \cdot \\ H_K(\omega) \end{bmatrix} = \sum_{\ell=0}^{M} \begin{bmatrix} C_{1\ell}e^{-j\omega\Delta_{1\ell}} \\ C_{2\ell}e^{-j\omega\Delta_{2\ell}} \\ \cdot \\ \cdot \\ \cdot \\ C_{K\ell}e^{-j\omega\Delta_{K\ell}} \end{bmatrix} \tag{5.5-2}$$

For the adaptive array processor the tap spacings are fixed throughout the training period but the weights are adjusted according to

$$\underline{W}_{j+1} = \underline{W}_j - 2\gamma_j\, \eta_j\, z_j + 2\gamma_j\, \underline{R}_{d\xi} \tag{5.5-3}$$

where γ_j determines the pitch of the algorithm and generally depends on the time index j ; the η's are the delayed inputs and z is the summer output. The variation of the weights during the training period will thus determine the adaptive behavior of the processor. The filtering problem was studied in the last two chapters where the variations of mean squared error as a function of time index j and input statistics were expressed explicitly. Here we shall study the detection problem by examining the variations of the performance criteria during the training period. Since the input is random, the weights expressed by (5.5-3) are random and only their expected values are of significance. It has been shown in Sect. 3.5, especially Eqs. (3.5-7), (3.5-35) and (3.5-43), that the expected values of $\underline{W}_j$ at any stage is related to their initial and final values by the general formula

$$E[\underline{W}_{j+1}] = p_j\, \underline{W}_1 + q_j\, \underline{W}_\infty \tag{5.5-4}$$

where $\underline{W}_1 = \underline{W}(j=1)$ and $\underline{W}_\infty = \underline{W}(j=\infty)$; p_j and q_j are functions of j and depend on the choice of the weighting sequence γ_j . For example, if γ_j is chosen as a weighting matrix such that

$$\gamma_j = \frac{1}{2(j+1)} \begin{bmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_{K(M+1)}} \end{bmatrix} \tag{5.5-5}$$

$$\lambda_k = k^{th} \text{ eigenvalue of } \underline{R}_\eta \,, \; i = 1,2,\cdots,K(M+1)$$

then

$$p_j = \frac{1}{j+1} \tag{5.5-6}$$

$$q_j = \frac{1}{j+1} \tag{5.5-7}$$

If, however

$$\gamma_j = \frac{1}{2(j+1)} \underline{I}, \quad \underline{I} = \text{unity matrix} \tag{5.5-8}$$

then they are of the form

$$\frac{\Gamma(j+2-\lambda_{max})}{\Gamma(2-\lambda_{max})(j+1)!} \le p_j \le \frac{\Gamma(j+2-\lambda_{min})}{\Gamma(2-\lambda_{min})(j+1)!} \tag{5.5-9}$$

$$q_j = 1 - p_j \tag{5.5-10}$$

where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of the input correlation matrix $\underline{R}_\eta$.

Combining Eqs. (5.5-2) and (5.5-4) we can write

$$\begin{aligned} \underline{H}_{j+1} &= p_j\, \underline{H}_1(\omega) + q_j\, \underline{H}_\infty(\omega) \\ &= p_j\, \underline{a}^* + q_j\, \underline{\Phi}_{nn}^{*-1}\, \phi_d\, \underline{a}^* \end{aligned} \tag{5.5-11}$$

$\underline{H}_1(\omega)$ and $\underline{H}_\infty(\omega)$ denote respectively the initial and final forms of transfer function vectors.

Since in all cases $p_o = 1$, $q_o = 0$ and $p_\infty = 0$, $q_\infty = 1$, the adaptive processor starts to be a square law detector and will be transformed gradually into an optimal one. We can determine where the adaptive processor stands between these bounds during the adaptation period by substituting Eq. (5.5-11) into the expressions of various performance criteria.

Basically, we are required to evaluate the following three integrals

$$A_{j+1} = \frac{1}{2\pi} \int_0^{\omega_o} |G|^2 \, (p_j\, \underline{H}_1 + q_j\, \underline{H}_\infty)^T \, \underline{\Phi}_{ss} \, (p_j\, \underline{H}_1 + q_j\, \underline{H}_\infty)^* d\omega \tag{5.5-12}$$

$$B_{j+1} = \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4 \{ (p_j \underline{H}_1 + q_j \underline{H}_\infty)^T \underline{\Phi}_{nn}(p_j \underline{H}_1 + q_j \underline{H}_\infty)^* \}^2 d\omega \tag{5.5-13}$$

$$C_{j+1} = \frac{1}{2\pi} \int_{0}^{\omega_o} |G|^2 (p_j \hat{\underline{H}}_1 + q_j \hat{\underline{H}}_\infty)^T \underline{\Phi}_{xx} (p_j \hat{\underline{H}}_1 + q_j \hat{\underline{H}}_\infty) d\omega \tag{5.5-14}$$

where in Eq. (5.5-14) the H's have steering vector $\hat{\underline{a}}$ rather the signal delays $\underline{a}$ in deriving the directivity patterns.

Since

$$(p_j \underline{H}_1 + q_j \underline{H}_\infty)^T \underline{\Phi} (p_j \underline{H}_1 + q_j \underline{H}_\infty)^*$$

$$= p_j^2 \underline{H}_1^T \underline{\Phi} \underline{H}_1^* + q_j^2 \underline{H}_\infty^T \underline{\Phi} \underline{H}_\infty^*$$

$$+ 2 p_j q_j \underline{H}_1^T \underline{\Phi} \underline{H}_\infty^* \tag{5.5-15}$$

we shall evaluate each one of the three integrals, Eqs. (5.5-12) through (5.5-14), by using Eq. (5.5-15). When the first two terms of Eq. (5.5-15) are used, the results are already available from the previous two sections on the initial and final behaviors.

Thus, we obtain $A_{j+1}^{(k)}$, k = 1,2,3, when the k^{th} term on the right-hand side of Eq. (5.5-15) is substituted into Eq. (5.5-12)

$$A_{j+1}^{(1)} = p_j^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2 \underline{H}_1^T \underline{\Phi}_{ss} \underline{H}_1^* = p_j^2 \bar{y}_{d.c}^{(1)} \tag{5.5-16}$$

$$A_{j+1}^{(2)} = q_j^2 \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2 \underline{H}_\infty^T \underline{\Phi}_{ss} \underline{H}_\infty^* = q_j^2 \bar{y}_{d.c}^{(\infty)} \tag{5.5-17}$$

where $\bar{y}_{d.c}^{(1)}$ is given by Eq. (5.3-3) and $\bar{y}_{d.c}^{(\infty)}$ by Eq. (5.4-19)

$$A_{j+1}^{(3)} = 2 p_j q_j \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2 \underline{H}_1^T \underline{\Phi}_{ss} \underline{H}_\infty d\omega$$

$$= 2 p_j q_j \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_d^{-1} \underline{a}^{*T} \phi_d \underline{a} \underline{a}^{*T} \underline{\Phi}_{nn}^{-1} \underline{a} \phi_d d\omega$$

which for broadside condition becomes

$$A_{j+1}^{(3)} = 2\, p_j\, q_j\, \frac{K}{2\pi} \int_0^{\omega_0} \frac{S}{N} \left[K - \frac{\sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(\rho_i - \rho_h)}}{K + N/I}\right] d\omega$$

$$= 2\, p_j\, q_j\, \frac{\omega_0 K}{2\pi} \left(\frac{S}{N}\right) \left\{ \frac{K-1+N/I}{K+N/I} \right.$$

$$\left. - \frac{2}{K+N/I} \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_0\, i\, \rho_0}{\omega_0\, i\, \rho_0} \right\} \qquad (5.5\text{-}18)$$

if $\omega_0 \rho_0 >> 1$, then

$$A_{j+1}^{(3)} \simeq p_j\, q_j\, \frac{K\omega_0}{\pi} \left(\frac{S}{N}\right) \frac{K-1+N/I}{K+N/I} \qquad (5.5\text{-}19)$$

We shall consider the second integral.

Note that

$$\left\{ (p_j\, \underline{H}_1 + q_j\, \underline{H}_\infty)^T\, \underline{\Phi}\, (p_j\, \underline{H}_1 + q_j\, \underline{H}_\infty)^* \right\}^2$$

$$= p_j^{\,4}\, (\underline{H}_1^T\, \underline{\Phi}\, \underline{H}_1^*)^2 + q_j^{\,4}\, (\underline{H}_\infty^T\, \underline{\Phi}\, \underline{H}_\infty^*)^2$$

$$+ 4\, p_j^{\,2}\, q_j^{\,2} (\underline{H}_1^T\, \underline{\Phi}\, \underline{H}_\infty^*)^2 + 4\, p_j^{\,2}\, q_j^{\,2} (\underline{H}_1^T\, \underline{\Phi}\, \underline{H}_1^*) (\underline{H}_\infty^T\, \underline{\Phi}\, \underline{H}_\infty^*)$$

$$+ 2\, p_j\, q_j^{\,3}\, (\underline{H}_\infty^T\, \underline{\Phi}\, \underline{H}_\infty^*)\, (\underline{H}_1^T\, \underline{\Phi}\, \underline{H}_\infty^*)$$

$$+ 2\, p_j^{\,3}\, q_j\, (\underline{H}_1^T\, \Phi\, \underline{H}_1^*)\, (\underline{H}_1^T\, \underline{\Phi}\, \underline{H}_\infty^*) \qquad (5.5\text{-}20)$$

Thus we obtain $B_{j+1}^{(k)}$, $k = 1,2,\cdots,6$, by substituting the k^{th} term on the right-hand side of Eq. (5.5-20) into Eq. (5.5-13)

$$B_{j+1}^{(1)} = p_j^{\,4}\, \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4\, (\underline{H}_1^T\, \underline{\Phi}_{nn}\, \underline{H}_1^*)^2\, d\omega$$

$$= p_j^{\,4}\, (\sigma_y^{(1)})^2 \qquad (5.5\text{-}21)$$

$$B_{j+1}^{(2)} = q_j^{\,4}\,\frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4 \,(\underline{H}_\infty^T\,\underline{\Phi}_{nn}\,\underline{H}_\infty^*)^2\, d\omega$$

$$= q_j^{\,4}\,(\sigma_y^{(\infty)})^2 \qquad (5.5\text{-}22)$$

where $(\sigma_y^{(1)})^2$ is given by Eq. (5.3-9) and $(\sigma_y^{(\infty)})^2$ by Eq. (5.4-22).

$$B_{j+1}^{(3)} = 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4\,(\underline{H}_1\,\underline{\Phi}_{nn}\,\underline{H}_\infty^*)^2\, d\omega$$

$$= 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{1}{\pi T_{av}} \int_{0}^{\omega_o} \phi_d^{-2}\,(\underline{a}^{*T}\,\underline{\Phi}_{nn}\,\underline{\Phi}_{nn}^{-1}\,\underline{a}\,\phi_d)^2\, d\omega$$

$$= 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{K^2\omega_o}{\pi T_{av}} \qquad (5.5\text{-}23)$$

$$B_{j+1}^{(4)} = 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4\,(\underline{H}_1^T\,\underline{\Phi}_{nn}\,\underline{H}_1^*)(\underline{H}_\infty^T\,\underline{\Phi}_{nn}\,\underline{H}_\infty^*)\, d\omega$$

$$= 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{1}{\pi T_{av}} \int_{0}^{\omega_o} \phi_d^{-2}\,(\underline{a}^{*T}\,\underline{\Phi}_{nn}\,\underline{a})(\phi_d^2\,\underline{a}^{*T}\,\underline{\Phi}_{nn}^{-1}\,\underline{a})\, d\omega$$

$$= 4\,p_j^{\,2}\,q_j^{\,2}\,\frac{1}{\pi T_{av}} \int_{0}^{\omega_o} (K + \frac{I}{N}\,|\underline{a}^{*T}\underline{b}|^2)(K - \frac{|\underline{a}^{*T}\underline{b}|^2}{K+N/I})\, d\omega \qquad 5.5\text{-}24)$$

which becomes for broadside condition

$$B_{j+1}^{(4)} = 4\,p_j^{\,2}\,q_j^{\,2} \Bigg\{ K^2 + \frac{K^2 I/N}{K+N/I}\,[K + \sum_{i=1}^{K-1} (K-i)\,\frac{\sin \omega_o\, i\, \rho_o}{\omega_o\, i\, \rho_o}]$$

$$+ \frac{I/N}{K+N/I}\,[K^2 + 4K \sum_{i=1}^{K-1} (K-i)\,\frac{\sin \omega_o\, i\, \rho_o}{\omega_o\, i\, \rho_o}$$

$$+ 2 \sum_{i=1}^{K-1} \sum_{h=1}^{K-1} (K-i)(K-h)\,(\frac{\sin \omega_o (i-h)\omega_o\rho_o}{\omega_o (i-h)\omega_o\rho_o} + \frac{\sin \omega_o (i+h)\rho_o}{\omega_o (i+h)\rho_o})] \Bigg\} \qquad (5.5\text{-}25)$$

Furthermore, if $\omega_o \rho_o >> 1$, then

$$B_{j+1}^{(4)} \simeq 4\, p_j^{\,2}\, q_j^{\,2}\, \frac{\omega_o}{\pi T_{av}} \left\{ K^2 + \frac{K^3 I/N}{K+N/I} + \frac{I/N}{K+N/I} [K^2 + 2 \sum_{i=1}^{K-1} (K-i)^2] \right\}$$

$$= 4\, p_j^{\,2}\, q_j^{\,2}\, \frac{\omega_o}{\pi T_{av}} \left\{ K^2 + \frac{I/N}{K+N/I} [K^3 + K^2 + \frac{1}{3} K(K-1)(2K-1)] \right\} \quad (5.5\text{-}26)$$

$$B_{j+1}^{(5)} = 2\, p_j\, q_j^{\,3}\, \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4 \,(\underline{H}_\infty^T\, \underline{\Phi}_{nn}\, \underline{H}_\infty^*)(\underline{H}_1^T\, \underline{\Phi}_{nn}\, \underline{H}_\infty^*)\, d\omega$$

$$= 2\, p_j\, q_j^{\,3}\, \frac{1}{\pi T_{av}} \int_0^{\omega_o} (\underline{a}^{*T}\, \underline{\Phi}_{nn}^{-1}\, \underline{a})(\underline{a}^{*T}\, \underline{\Phi}_{nn}\, \underline{\Phi}_{nn}^{-1}\, \phi_d\, \underline{a})\, d\omega$$

$$= 2\, p_j\, q_j^{\,3}\, \frac{K}{\pi T_{av}}\, \frac{S}{N} \int_0^{\omega_o} (K - \frac{|\underline{a}^{*T}\underline{b}|^2}{K+N/I})\, d\omega \quad (5.5\text{-}27)$$

for $\tau_i = 0$, the above expression reduces to

$$B_{j+1}^{(5)} = 2\, p_j\, q_j^{\,3}\, \frac{K\omega_o}{\pi T_{av}} (\frac{S}{N}) [\frac{K(K-1+N/I)}{(K+N/I)} - 2 \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o\, i\, \rho_o}{\omega_o\, i\, \rho_o}]$$

$$\simeq 2\, p_j\, q_j^{\,3}\, (\frac{S}{N})\, \frac{\omega_o}{\pi T_{av}}\, K^2\, \frac{K-1+N/I}{K+N/I} \quad \text{for } \omega_o \rho_o >> 1 \quad (5.5\text{-}28)$$

$$B_{j+1}^{(6)} = 2\, p_j^{\,3}\, q_j\, \frac{1}{\pi T_{av}} \int_{-\infty}^{\infty} |G|^4 \,(\underline{H}_1^T\, \underline{\Phi}_{nn}\, \underline{H}_1^*)(\underline{H}_1^T\, \underline{\Phi}_{nn}\, \underline{H}_\infty^*)\, d\omega$$

$$= 2\, p_j^{\,3}\, \frac{1}{\pi T_{av}} \int_0^{\omega_o} \phi_d^{-2}\, (\underline{a}^{*T}\, \underline{\Phi}_{nn}\, \underline{a})(\underline{a}^{*T}\, \underline{a}\, \phi_d)\, d\omega$$

$$= 2\, p_j^{\,3}\, q_j\, \frac{K}{\pi T_{av}} \int_0^{\omega_o} \frac{N}{S} (K + \frac{I}{N} |\underline{a}^{*T}\, \underline{b}|^2)\, d\omega$$

$$= 2\, p_j^{\,3}\, q_j\, \frac{K\omega_o}{\pi T_{av}} (\frac{N}{S}) \left\{ K + \frac{I}{N} [K + 2 \sum_{i=1}^{K-1} (K-i) \frac{\sin \omega_o\, i\, \rho_o}{\omega_o\, i\, \rho_o}] \right\}$$

for $\tau_i = 0$ and

$$= 2\, p_j^{\,3}\, q_j \frac{K^2 \omega_o}{\pi T_{av}} \left(\frac{N}{S}\right)\left(1 + \frac{I}{N}\right) \tag{5.5-29}$$

for $\rho_o \omega_o >> 1$.

We shall next consider the third integral. Using the k^{th} term on the right-hand side of Eq. (5.5-14) in evaluating Eq. (5.5-13), we have $c_{j+1}^{(k)}$, $k = 1,2,3$

$$c_{j+1}^{(1)} = p_j^{\,2} \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{\hat{H}}_1^T\, \underline{\Phi}_{xx}\, \underline{H}_1^*\, d\omega = p_j^{\,2}\, \overline{y}_1 \tag{5.5-30}$$

$$c_{j+1}^{(2)} = q_j^{\,2} \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{\hat{H}}_\infty^T\, \underline{\Phi}_{xx}\, \underline{\hat{H}}_\infty^*\, d\omega = q_j^{\,2}\, \overline{y}_\infty \tag{5.5-31}$$

where $\overline{y}_1$ is given by Eq. (5.3-22) and $\overline{y}_\infty$ by Eq. (5.4-32).

$$c_{j+1}^{(3)} = 2\, p_j\, q_j \frac{1}{2\pi} \int_{-\infty}^{\infty} |G|^2\, \underline{\hat{H}}_1^T\, \underline{\Phi}_{xx}\, \underline{\hat{H}}_\infty\, d\omega$$

$$= \frac{p_j q_j}{\pi} \int_0^{\omega_o} \underline{\hat{a}}^{*T} \left(\phi_d\, \underline{a}\, \underline{a}^{*T} + \underline{\Phi}_{nn}\right) \underline{\Phi}_{nn}^{-1}\, \underline{\hat{a}}\, d\omega$$

$$= \frac{p_j q_j}{\pi} \int_0^{\omega_o} \left\{ K + \frac{S}{N} \left[|\underline{\hat{a}}^{*T}\, \underline{a}|^2 - \frac{1}{K+N/I}\, \underline{\hat{a}}^{*T}\, \underline{a}\, \underline{a}^{*T}\, \underline{b}\, \underline{b}^{*T}\, \underline{\hat{a}} \right] \right\} d\omega \tag{5.5-32}$$

Since

$$|\underline{\hat{a}}^{*T}\, \underline{a}|^2 = \sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(i-h)(\tau_o - \hat{\tau}_o)} \tag{5.5-33}$$

and

$$\underline{\hat{a}}^{*T}\, \underline{a}\, \underline{a}^{*T}\, \underline{b}\, \underline{b}^{*T}\, \underline{\hat{a}}$$

$$= \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{k=1}^{K} e^{j\omega(\tau_i - \hat{\tau}_i - \tau_h + \rho_h - \rho_k - \hat{\tau}_k)}$$

$$= K \sum_{i=1}^{K} \sum_{h=1}^{K} e^{j\omega(i-h)(\tau_o - \hat{\tau}_o)}$$

$$+ \sum_{i=1}^{K} \sum_{k=1}^{K} e^{j\omega(i-k)(\rho_o - \hat{\tau}_o)}$$

$$+ K \sum_{h=1}^{K} \sum_{k=1}^{K} e^{j\omega(k-h)(\rho_o - \tau_o)}$$

$$- 2K + \sum_{i=1}^{K} \sum_{h=1}^{K} \sum_{k=1}^{K} e^{j\omega(\tau_i - \hat{\tau}_i - \tau_h + \rho_h - \rho_k - \hat{\tau}_k)}$$

$$i \neq h \neq k \qquad (5.5\text{-}34)$$

Eq. (5.5-32) reduces to the following expression by omitting the contribution due to the last term on the right-hand side of Eq. (5.5-34)

$$C_{j+1}^{(3)} = p_j q_j \frac{\omega_o}{\pi} [K + \frac{S}{N} K - \frac{S}{N} \frac{1}{K+N/I} (3K^2 - 2K)]$$

$$+ 2 p_j q_j \frac{\omega_o}{\pi} \frac{S}{N} \sum_{i=1}^{K-1} (K-i) \Big\{ \frac{\sin \omega_o i(\tau_o - \hat{\tau}_o)}{\omega_o i(\tau_o - \hat{\tau}_o)} (1 + \frac{K}{K+N/I})$$

$$\frac{K}{K+N/I} [\frac{\sin \omega_o i(\rho_o - \hat{\tau}_o)}{\omega_o i (\rho_o - \hat{\tau}_o)} + \frac{\sin \omega_o i(\rho_o - \tau_o)}{\omega_o i(\rho_o - \tau_o)}] \Big\} \qquad (5.5\text{-}35)$$

Just as we did in expressions $\overline{y}_1$ and $\overline{y}_\infty$ for three special cases, we shall also evaluate $C_{j+1}^{(3)}$ in a similar fashion. When the array is steering in a direction far away from that of target and interference, we neglect all the oscillatory terms and get

$$C_{j+1}^{(3)}(\theta) = p_j q_j \frac{\omega_o}{\pi} K[1 + \frac{S}{N} - \frac{S}{N} \frac{1}{K+N/I} (3K - 2)] \qquad (5.5\text{-}36)$$

In the target direction, $\hat{\tau}_o = \tau_o$

$$C_{j+1}^{(3)}(\theta = \theta_T) = p_j q_j \frac{\omega_o}{\pi} [K + \frac{S}{N} K - \frac{S}{N} \frac{1}{K+N/I} (3K^2 - 2K)]$$

$$+ 2p_j q_j (\frac{S}{N}) \frac{\omega_o}{\pi} \frac{2K+N/I}{K+N/I} \sum_{i=1}^{K-1} (K-i)$$

$$= p_j q_j \frac{\omega_o}{\pi} K[1 + \frac{S}{N} \frac{2K^2-4K+2+KN/I}{K+N/I}] \tag{5.5-37}$$

and in the interference direction

$$C^{(3)}_{j+1} (\theta = \theta_I) \simeq p_j q_j \frac{\omega_o}{\pi} [K + \frac{S}{N} K - \frac{S}{N} \frac{1}{K+N/I} (3K^2 - 2K)]$$

$$+ 2 p_j q_j (\frac{S}{N}) \frac{\omega_o}{\pi} \frac{K}{K+N/I} \sum_{i=1}^{K-1} (K-i)$$

$$= p_j q_j \frac{\omega_o}{\pi} K[1 + \frac{S}{N} \frac{K^2- 3K+2+N/I}{K+N/I}] \tag{5.5-38}$$

Now we are ready to express the performance variations during the training period.

(1) Output signal-to-noise ratio

$$SNR_{j+1} = \frac{\sum_{k=1}^{3} A^{(k)}_{j+1}}{\left(\sum_{k=1}^{6} B^{(k)}_{j+1} \right)^{1/2}} \tag{5.5-39}$$

where the A's are given by Eqs. (5.5-16) through (5.5-19) and the B's by Eqs. (5.5-21) through (5.5-29)

(2) Directivity Pattern

$$\bar{y}_{j+1} = \sum_{k=1}^{3} C^{(k)}_{j+1} \tag{5.5-40}$$

where the C's are given by Eqs. (5.5-30), (5.5-31) and (5.5-35). For any particular steering direction $\bar{y}_{j+1}$ is readily computed from Eqs. (5.3-23) through (5.3-25), (5.4-33) through (5.3-35), and (5.5-36) through (5.5-38). Various computations are shown in Chapter Six.

CHAPTER SIX

COMPUTER SIMULATIONS AND NUMERICAL EXAMPLES

6.1 Introduction

A great deal of attention has been given to proving that iterative procedures described in previous chapters converge under certain conditions. Having proved convergence of the adjustment schemes, our problem is to demonstrate that the procedures are feasible; that is, solutions can be obtained by using the adjustment procedures in a reasonable amount of time. To establish this point, computer studies were made to the design of adaptive tapped-delay-line filters and detectors. Recall that the approach to adaptive receiver design has been an optimal one. The adaptive design is a result of realizing the optimum receiver in a sequential manner.

In this chapter we consider simulating an adaptive processor on a digital computer for a rather specific case to observe how the processor performances vary with time. Several examples have been worked out by digital computer simulations to verify all the theoretical analyses presented in Chapter III.

Some computations have also been carried out to show the performances of an adaptive detector described in Chapter V. These computations were done based on theoretical analysis rather than simulation which, in this case, would require too much computing time without providing any general conclusions. All these numerical examples were worked out on the IBM 7094 II-7040 direct-coupled system at the Yale Computer Center.

6.2 Computer Simulations

An arbitrary array processor was used here as an example to demonstrate the properties of the two algorithms given in (3.3-6) and (3.3-13), i.e., the algorithms using desired signal and signal correlation function, respectively.

A linear array of six uniformly spaced isotropic hydrophones was assumed to

be influenced by the following set of signals:

1. A planar target waveform incident from the broadside.
2. A single interfering planar noise waveform incident at angle $\theta_I = 120^o$.
3. A "white" gaussian noise source at each hydrophone representing the ambient noise which is assumed to be uncorrelated from hydrophone to hydrophone.

The hydrophones were spaced c/ω_0 units apart, where c is the velocity of propagation in the isotropic medium and ω_0 is the center frequency of the target signal. The output of each sensor was processed using a tapped-delay line containing ten multiplying weights and nine ideal time delay of $(\frac{1}{2\omega_0})$ seconds each. Because the target-signal waveform was incident from the broadside direction, the target signal arrived simultaneously -- i.e., in phase -- at the output of all six hydrophones.

The target signal and the interference were modeled as a broadband gaussian random processes. At each hydrophone the ratio of target-signal variance to total noise variance is 0.01. All these properties were generated by passing a pseudo-random gaussian sequence through an appropriately designed digital filter. Signal, noise, and interferences were generated as sequences of random numbers from random number generators. The sequences were then transformed from rectangular distribution to normal distribution using existing programs. Each simulation started with an initial weight vector having all components associated with no-delays set to unity and the rest to zero. The weight vector was then adapted using the appropriate iterative equations.

Throughout the study the sequence γ_j was determined as $\gamma_j = \frac{\gamma_o}{2(j+1)}$, γ_o being a variable parameter. The behavior of the process depends critically on the parameter γ_o ; hence each case considered was carried out for a number

of values of γ_o. As γ_o was increased, the convergence of the process increased steadily until a point was reached for which the process would break into violent oscillations that took a long time to die out. This point was predicted in Chapter III that γ_j should satisfy

$$0 < \gamma_j < \frac{1}{4\lambda_{max}}$$

for all j, especially at the starting moments. Since the largest eigenvalue λ_{max} was not known a priori, the best weighting sequence γ_j can only be chosen by experiments.

As a check case, the optimum (mean-square sense) values of the coefficients were found by correlation techniques, by averaging the necessary values of the correlation functions $R_s(\tau)$ and $R_n(\tau)$ over an interval of 2000 samples. This set of coefficients is compared in Table 1 with the sets of coefficients obtained by the approximation method for the two algorithms chosen. The average filtered error as measured by the algorithms over the 2000-sample interval is plotted against time index during the adaptation period. This is shown in Fig. 8 where the minimum mean squared error with the optimum weight is denoted by a horizontal line. The smooth curve indicates the mean squared error calculated by theoretical analysis.

Fig. 9 shows how the weight vector approaches its optimum point independent of the initial settings.

Figs. 10a and 10b show that faster rate of convergence can be obtained by increasing γ_o if $\gamma_j = \frac{\gamma_o}{2(j+1)}$, but not too large to cause oscillation. When constant weighting sequence was used (Fig. 10a), the mean squared error at later adaptation stages would oscillate around the minimum mse instead of approaching it gradually. For a single filter and known correlation functions, the relationships between the rate of convergence and the maximum eigenvalue are shown in

Fig. 10b. As evidenced by Table 1 and Figs. 8 and 10, the filter designs for the two algorithms are equivalent in the sense that they result in nearly equal values of average filtered error. The average filtering errors as obtained by these two algorithms over the 2000-sample interval were only a few percent greater than that for the filter designed by correlation techniques. In view of the limitation on the length of data available, this performance is entirely satisfactory.

The important point to be brought out is that the total computing time required by any one of the two adjustment procedures for finding the optimum set of coefficients was no greater than the computing time required to measure the necessary correlation functions and solve the associated set of simultaneous equations for the minimum mean-square error coefficients. Thus the adjustment methods are no more trouble to apply than correlation techniques, yet they eliminated the requirement of a priori statistics.

The effect of uncertain signal is shown in Fig. 11. We see that if the assumed signal power differs from the actual power by a multiplicative constant, the gains adjusted according to algorithm (3.3-16) will converge to their optimum values muitiplied by the same constant.

An attempt was made to compare the rate of convergence for two different approaches: the Kalman filtering technique using all the past information (see Section 4.3) versus the ordinary method of stochastic approximation. As expected, the Kalman technique gives a faster rate of convergence. Some of the reasons were given in Section 4.3. See Fig. 12.

There is no general method to select the right number of taps so that a predetermined accuracy can be achieved for any given system. Several runs were made to plot the minimum mse versus the number of taps. It was found that by properly adjusting the tap spacings, five or six taps could produce quite satisfactory results. One plot is shown in Fig. 13.

Coef. j		C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}	C_{19}
Start j = 1		1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Final j = 2000	(a)	0.3015	0.272	0.106	0.053	0.022	0.010	0.004	-0.003	0.004	-0.003
	(b)	0.3021	0.264	0.105	0.047	0.027	0.008	-0.003	0.001	0.003	-0.001
Opt. Values		0.3015	0.287	0.103	0.059	0.033	0.019	0.009	0.002	0.001	0.001

Table 1. Comparison of coefficients for the first individual filter.

(a) desired signal used

(b) signal correlation used

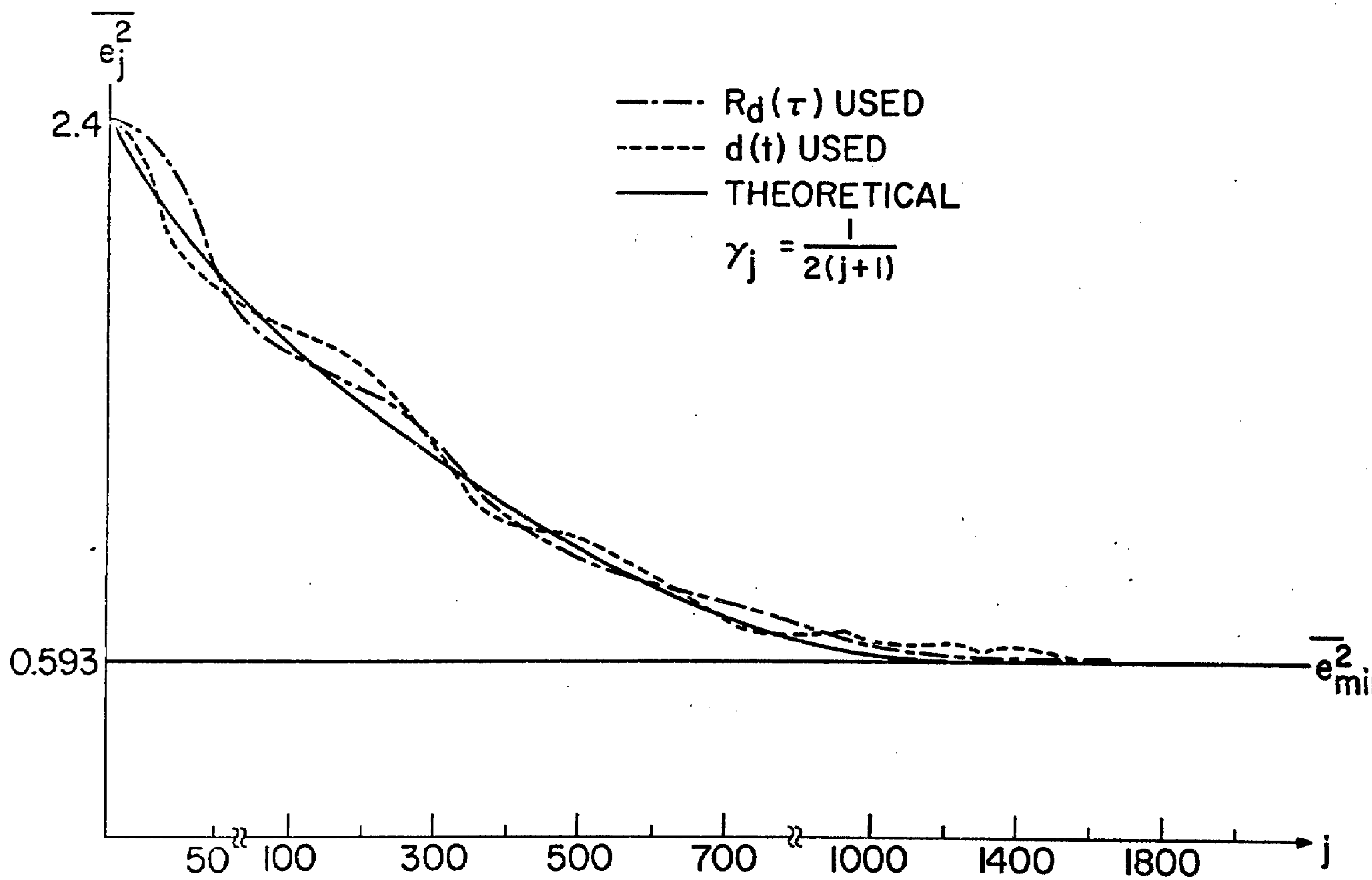


Figure 8. Variation of mean-squared error.

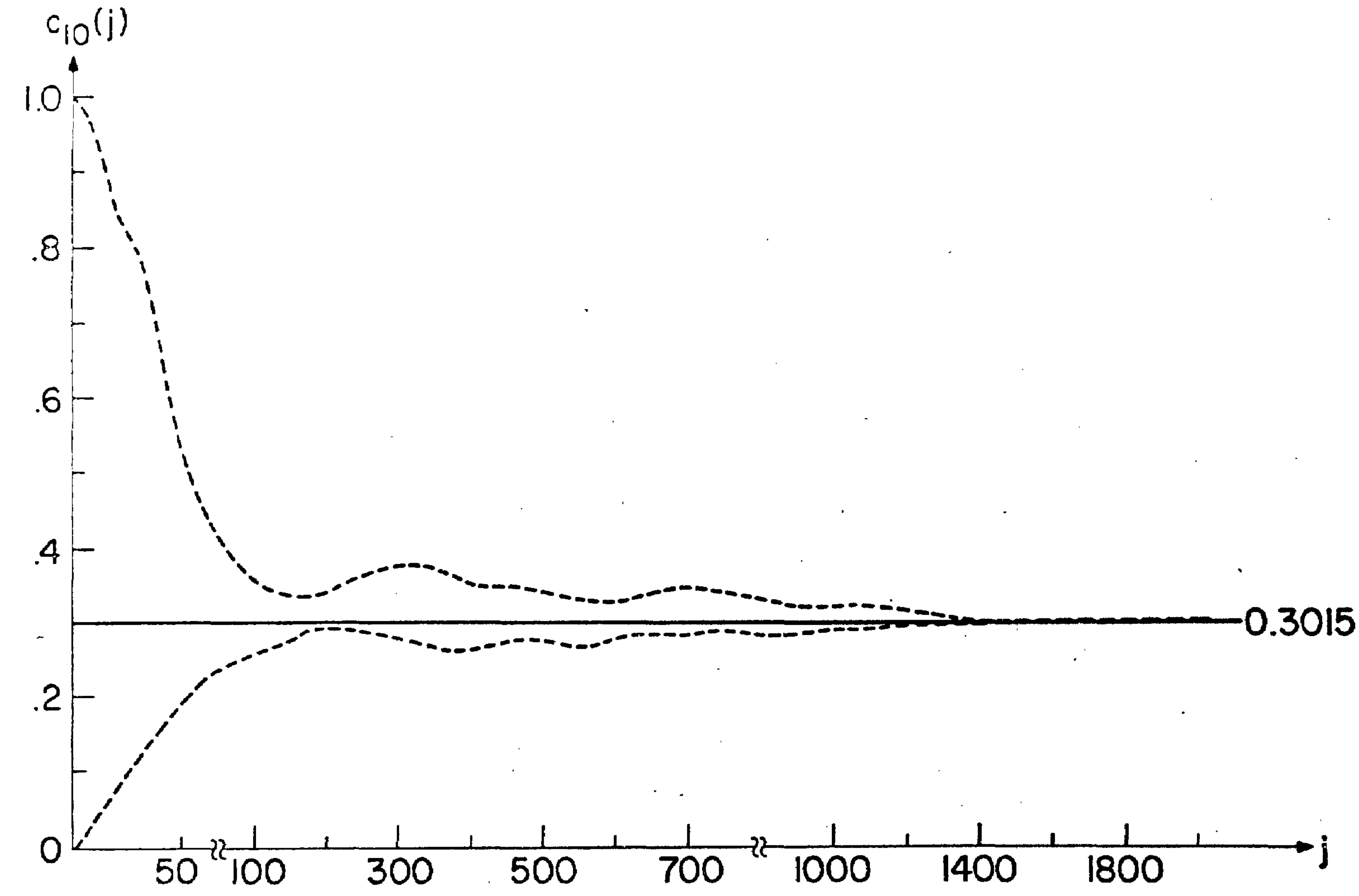


Figure 9. Variation of filter coefficients

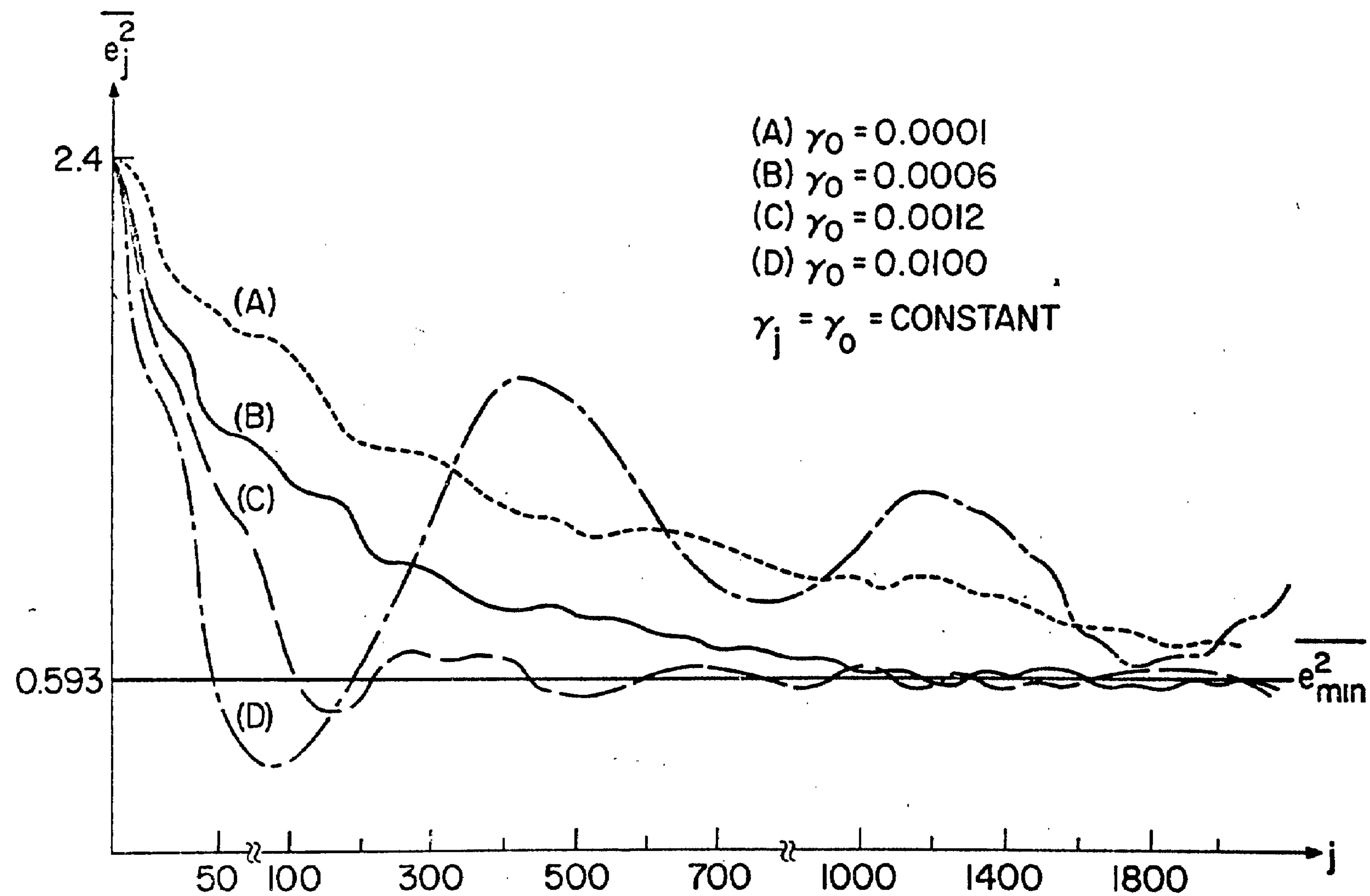


Figure 10a. MSE variation vs. weighting constant

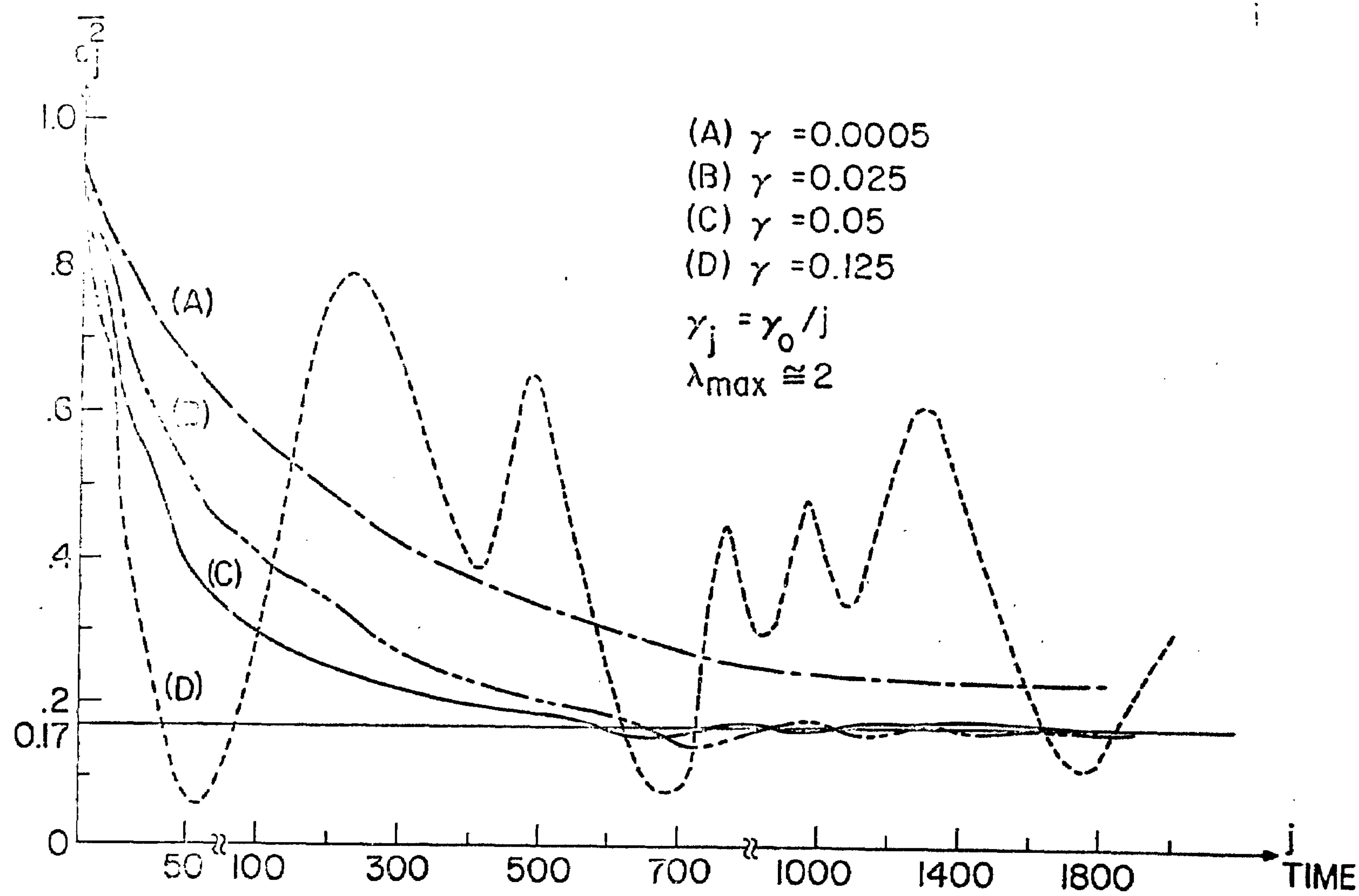


Figure 10a. MSE variation vs. weighting sequence
(single filter, $R(T) = R_o e^{a|T|}$, $R_o = 0.68$, $a \cong 0.67$

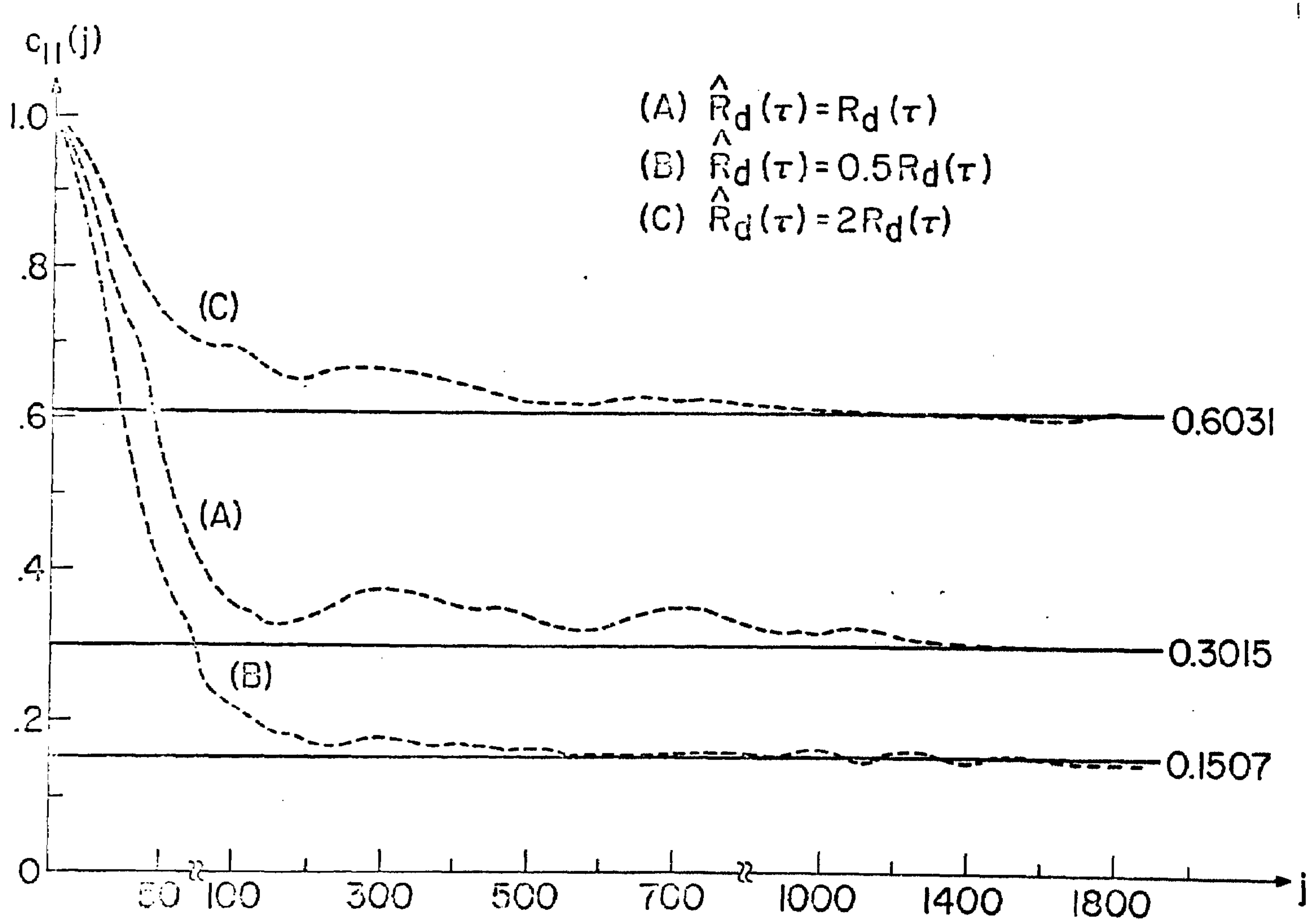


Figure 11. Effect of uncertain signal power

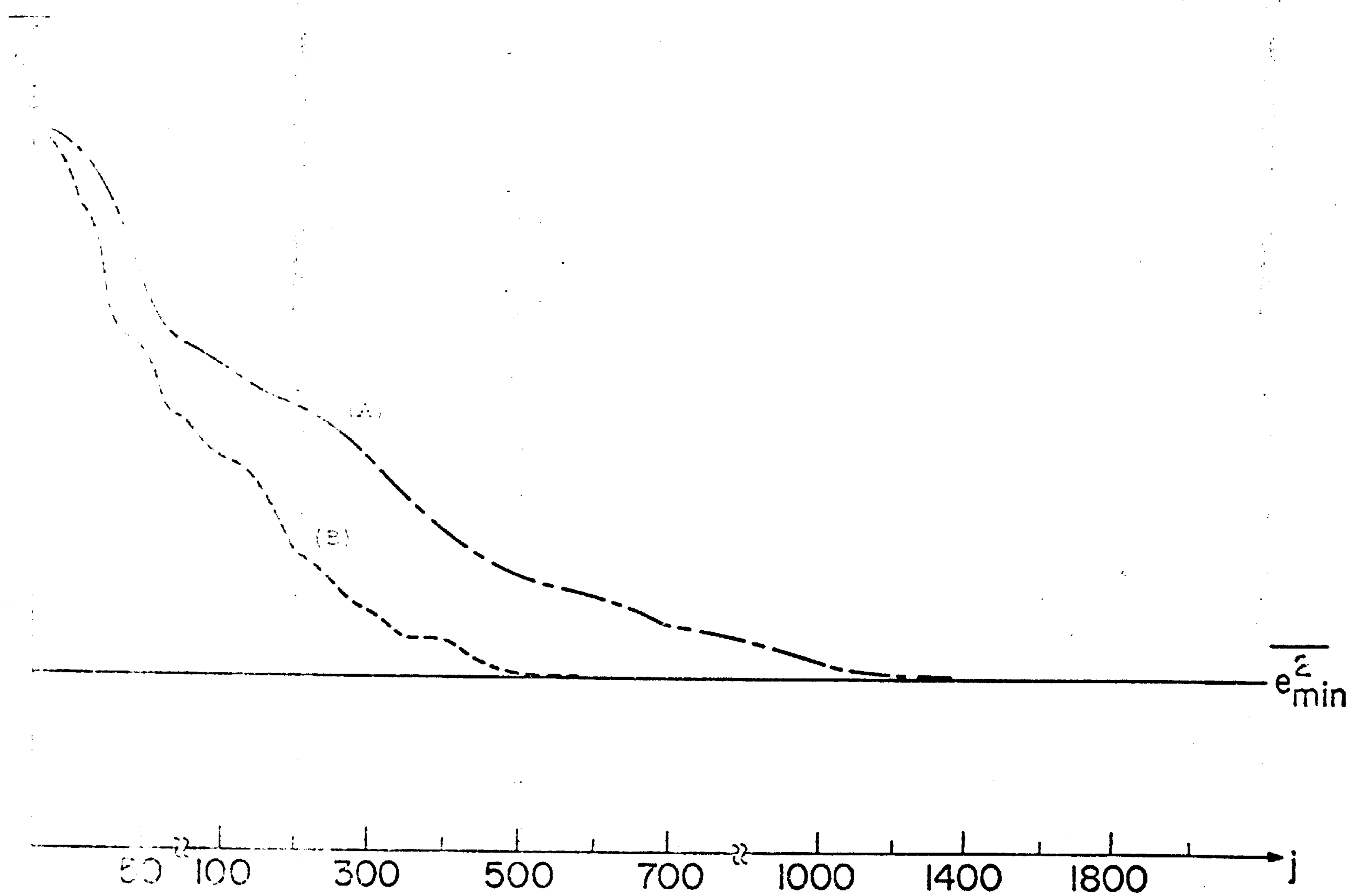


Figure 12. Comparison of rates of convergence
(A) Stochastic approximation
(B) Kalman filtering technique

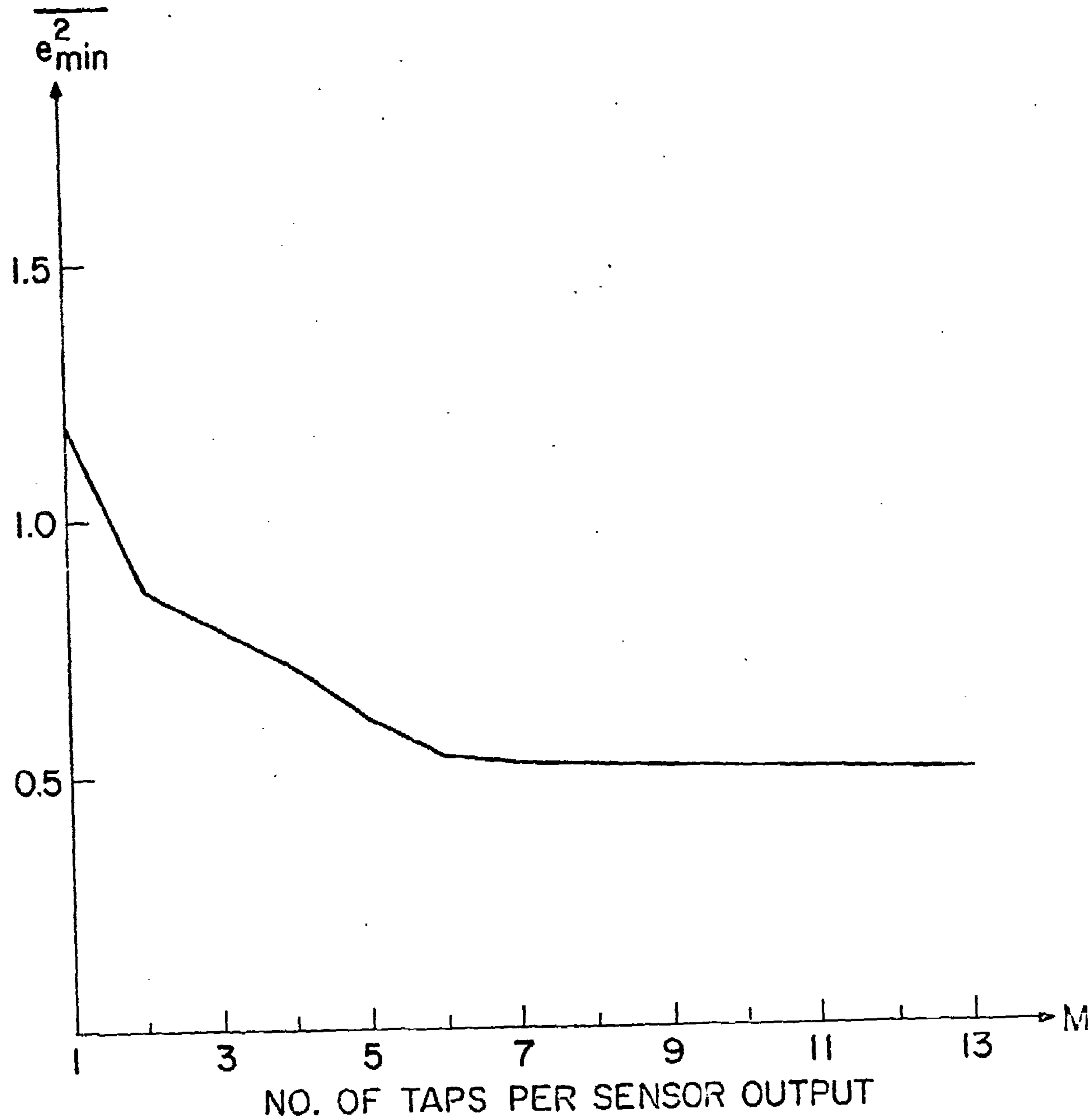


Figure 13. Minimum MSQ vs. number of taps per hydrophone output

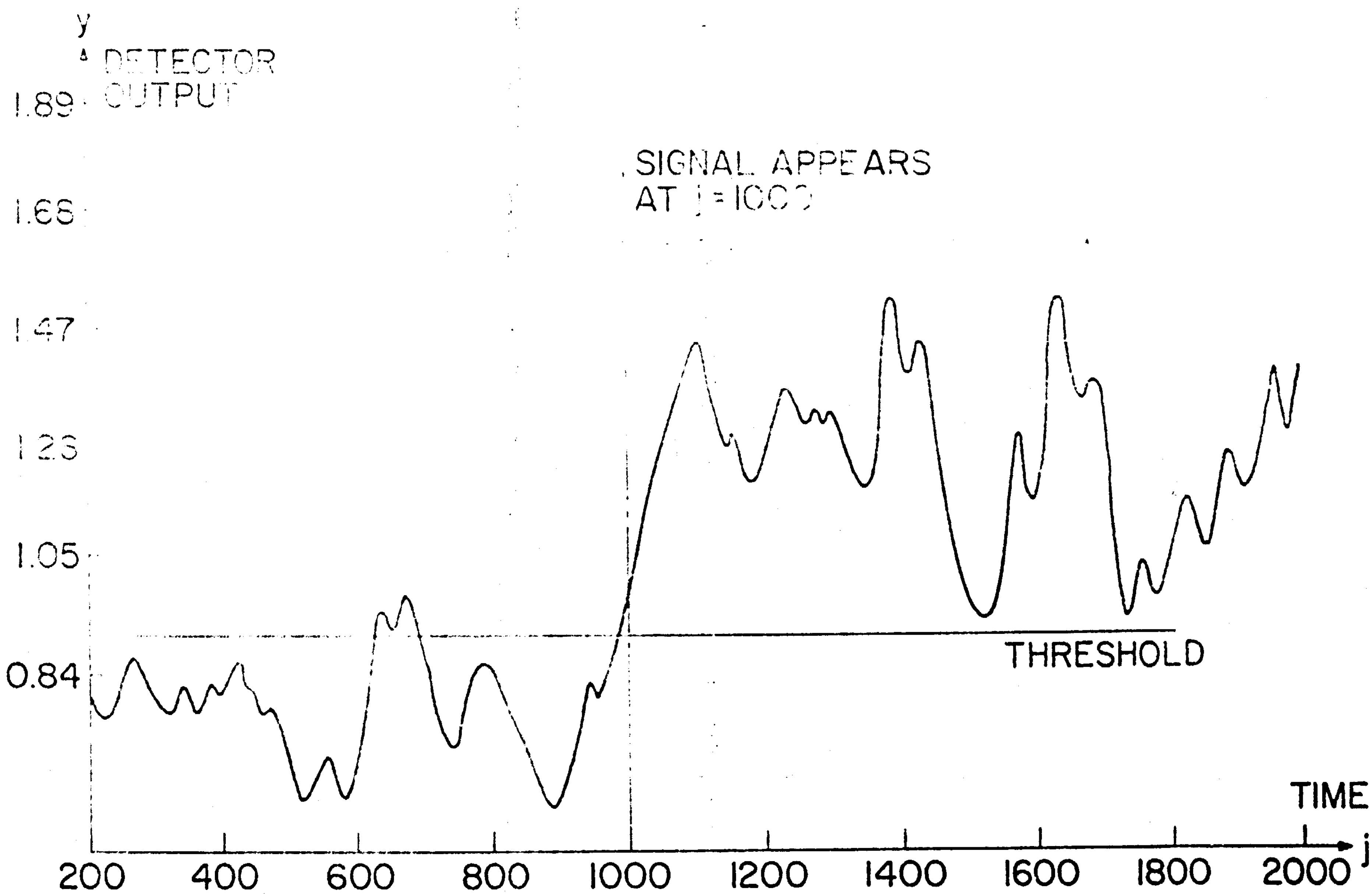


Figure 14. Variation of detector output

When this simulated array processor was used as a detector, the detector output was examined. After 2000 samples of adaptation, the weight vector was held fixed and the same 2000 samples were sent to the detector. The input contained noise only at the beginning of this operation and contained both signal and noise starting from $t = 1000$. Fig. 14 shows clearly how to interpret the detector output and decide the presence of a target.

6.3 Experimental Results

Sonar noise recorded at sea from a collection of 6 hydrophones has been used to test the iterative rules described in previous sections. An IBM 1800 computer was used to make data tapes compatible with an IBM 7094-7040 system which was used as the principal computational tool in the experiments.

The noise was a 2-second noise sampled every 1/8000 second. The total bandwidth of the data was about 425 Hz to 2400 Hz. The hydrophones in a linear array were separated by 7.5 inches.

Since the data were collected in actual sea water, the directionality of the noise field was not known exactly. From the display of correlation functions of several channels (See Fig. 15) there seemed to be some interfering source present in addition to the ambient noise.

In order to show how the processor eliminates background noise, a target signal was produced from a noise generator and passed through a filter. The signal autocorrelation function is shown in Fig. 16. Three different signal directions were tested, i.e., direction A (opposite to the assumed noise direction, direction B (perpendicular to the assumed noise direction), direction C (similar to the assumed noise direction). After proper signal delays we obtained three different noise correlation matrices whose cross-correlation coefficients for the six channels are tabulated in Table 2.

It is seen from Table 2 that for direction B the actual noise correlation matrix consists of many oscillatory terms. If signal delays were inserted to

align target coming from direction C, the interference was also in phase so that all the cross-correlation coefficients are positive. On the other hand, if signal delays were inserted to align the target coming from direction A, the noise field was further de-correlated. Consequently, the array gain defined by

$$\frac{\text{signal-to-noise ratio of the summer output}}{\text{signal-to-noise ratio at channel 1}}$$

was largest if the target came from direction A and least if both the target and the noise came from the same direction. This is shown in the second row of Table 3. Note that the signal-to-noise ratio in this experiment has been defined as the signal power divided by the noise power, rather than the d.c. change of the output due to the presence of the target divided by the rms fluctuation of the output. The later definition is more meaningful for the detection problem and the former definition is useful for the problem of signal extraction. Several cases were studied on how our proposed adaptive array processor eliminated the undesired noise. Complete results are shown in Table 3, where SNR_{in} is the input signal-to-noise ratio at Channel 1, SNR_c is the output signal-to-noise ratio of the conventional processor, SNR_6 is the final output signal-to-noise ratio of the adaptive processor after 2000 iterative adjustments and using 6 taps in each individual filter, SNR_{12} is the same as SNR_6 except that twelve taps were used for each individual filter. So far we have assumed that the reference or desired signal d(t) is the same as the target signal, i.e., d(t) = s(t). If d(t) is replaced by some delayed version of s(t), i.e.,

$$d(t) = s(t - \tau)$$

then, for some proper choice of τ, smaller mean-squared error or larger signal-to-noise ratios may result. Further discussions on this point can be found in [1]. For $\tau \neq 0$, the corresponding SNR_6 and SNR_{12} are denoted by SNR'_6 and SNR'_{12}. The improvement of SNR produced by the adaptive processors over that

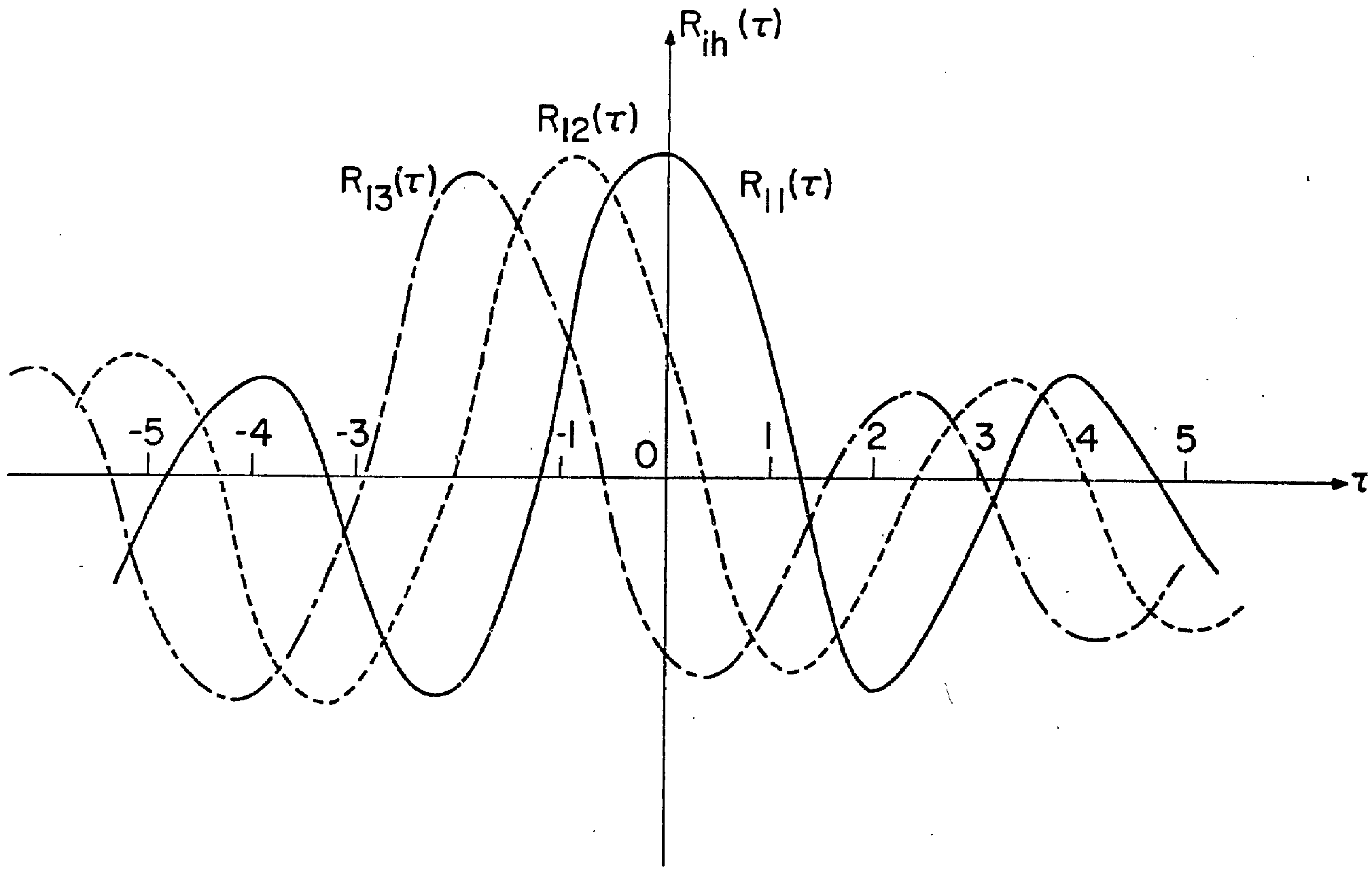


Figure 15. Noise correlation functions

(a) Channel	1	2	3	4	5	6
1	1.00	0.70	0.55	0.56	0.43	0.24
2	0.70	1.00	0.64	0.58	0.40	0.24
3	0.55	0.64	1.00	0.68	0.44	0.21
4	0.56	0.58	0.68	1.00	0.58	0.26
5	0.43	0.40	0.44	0.58	1.00	0.34
6	0.24	0.24	0.21	0.26	0.34	1.00

(b) Channel	1	2	3	4	5	6
1	1.00	0.42	-0.27	-0.13	0.10	0.089
2	0.42	1.00	0.36	-0.34	0.05	0.11
3	-0.27	0.36	1.00	0.22	-0.23	0.10
4	-0.13	-0.34	0.22	1.00	0.29	-0.13
5	0.10	0.05	-0.23	0.29	1.00	0.27
6	-0.089	0.11	0.10	-0.13	0.27	1.00

(c) Channel	1	2	3	4	5	6
1	1.00	-0.36	0.21	-0.13	0.12	-0.11
2	-0.36	1.00	-0.29	0.24	-0.21	0.07
3	0.21	-0.29	1.00	-0.33	0.15	-0.13
4	-0.13	0.24	-0.33	1.00	-0.24	0.12
5	0.12	-0.21	0.15	-0.24	1.00	-0.07
6	-0.11	0.07	-0.13	0.12	-0.07	1.00

TABLE [illegible] Noise cross-correlation coefficients for signal coming from (a) Direction C, (b) Direction B, (c) Direction A.

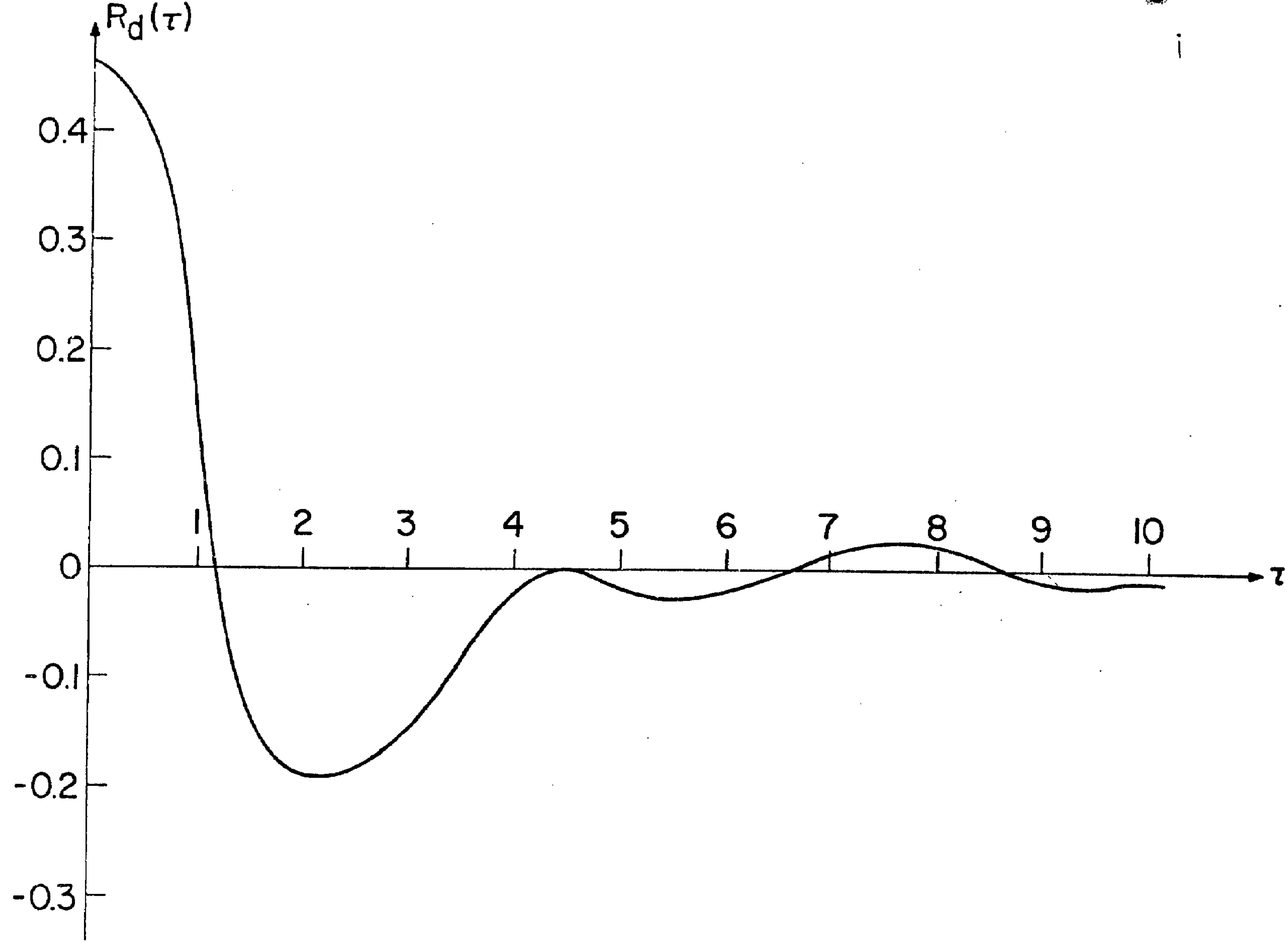


Figure 16. Signal correlation function

Items \ Cases	Direction C	Direction B	Direction A
SNR_c	0.002431	0.006676	0.012054
$\frac{SNR_c}{SNR_{IN}}$	2.38	6.55	11.8
$SNRO_6$	0.003967	0.009556	0.02594
$SNRO'_6$	0.004135	0.010772	0.03179
$SNRO_{12}$	0.004001	0.009887	0.02774
$SNRO'_{12}$	0.004187	0.010955	0.03293
$\frac{SNRO_6}{SNR_c}$	2.13 db	1.56 db	3.18 db
$\frac{SNRO'_6}{SNR_c}$	2.31 db	2.06 db	4.24 db
$\frac{SNRO_{12}}{SNR_c}$	2.18 db	1.73 db	3.63 db
$\frac{SNRO'_{12}}{SNR_c}$	2.36 db	2.15 db	4.41 db

TABLE 3. Experimental results (SNR_{IN}=0.001039 for all cases)

by the conventional processor such as SNR'_6/SNR_c are measured in db. These results shown here are remarkably close to the optimal filtering using complete input statistics (performed independently by R. Kneipfer of U.S. Underwater Sound Laboratory, New London, Connecticut).

6.4 Numerical Computations

To investigate how an adaptive detector changes its performance during the training period, we could, in principle, simulate such a processor in digital computers. However, there exist some practical difficulties. Since detection performances (output SNR, directivity patterns) are functions of output mean and variance, at each stage of the training process we are required to calculate the output and variance using sufficiently large numbers of samples (say, 1000 or more) for $\underline{W}_j$, where $j = 1,2,\ldots,$ number of test samples. Furthermore, if we want to change any one of the many system parameters such as number of hydrophones, number of taps, or input statistics, the whole process would have to be repeated.

In light of the above difficulties, analytic expressions were derived in Chapter V to determine how the adaptive detector performs for a specific case in which the input spectra are identical over a certain frequency range. Equations (5.5-39) and (5.5-40) are used extensively to carry out numerical computations.

Fig. 17 shows the variation of output signal-to-noise ratios during the adaptation period.

If target and interference are well separated in bearing, the (normalized) directivity patterns are shown in Fig. 18. In Fig. 18 computations were made using approximate expressions in the target direction $(\theta = \theta_T)$, interference direction $(\theta = \theta_I)$, and remote from both. Optimal $(j = \infty)$ behaviors of the array processor as a function of other system parameters have been considered previously by Schultheiss [32] and are not plotted here.

Experimental results using sonar data have verified that practical adaptive array processors can perform nearly as well as optimum processors in a stationary environment. It should be possible to adopt similar iterative processors to seismic and electromagnetic arrays which operate in a directional noise environment. It might be possible to minimize reverberation as well as ambient noise in systems where reverberation is significant.

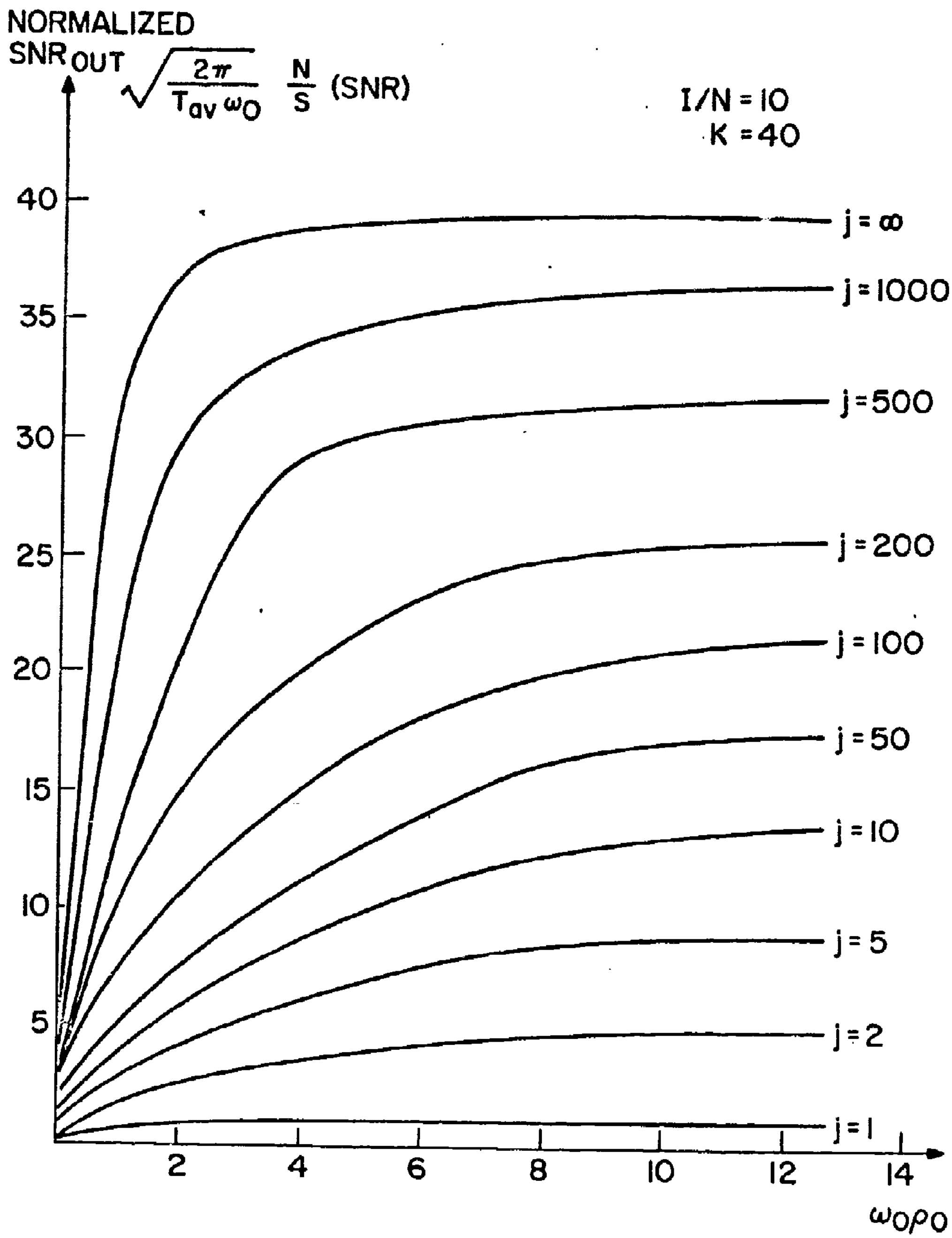


Figure 17. Variation of Output SNR

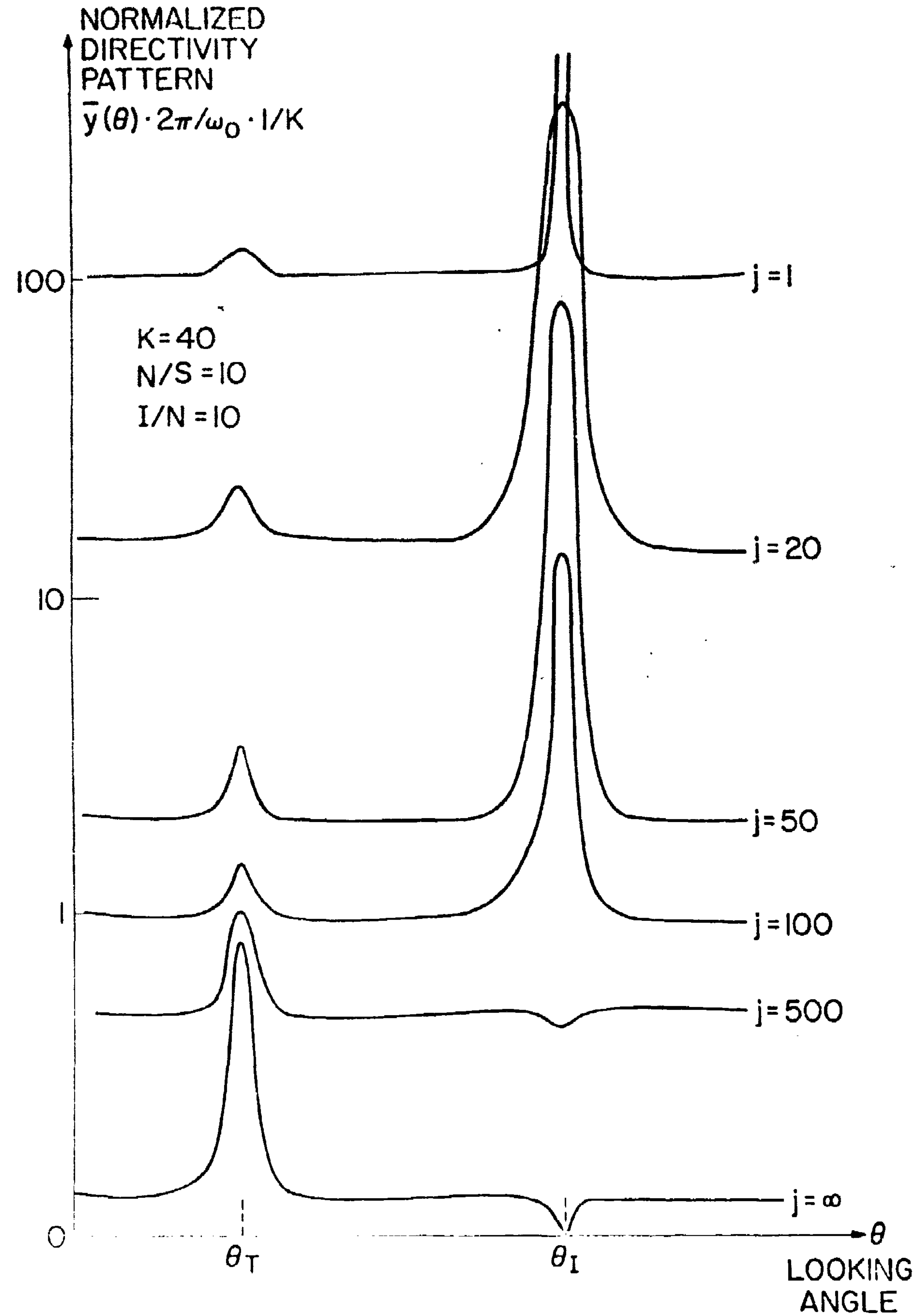


Figure 18. V... tion of Directivity Pattern

CHAPTER SEVEN

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH

7.1 Summary and Conclusion

The research described herein has developed a system for processing the outputs of a passive array of hydrophones. The system consists of an adaptive linear multichannel filter, together with algorithms for iterative adjustment of the weights on the topped-delay lines. It is designed to process the received wavefront in the presence of ambient noise and interferences. The system is designed in such a way that it can be readily implemented and be able to operate well in real time in the presence of noise fields whose statistics are unknown a priori.

a) Assumptions

The development and analysis of the array processor presented in this research has been based on the assumptions that

(1) Target, interferences and ambient noise are assumed to be gaussian random processes.

(2) The sum of interferences, ambient noise and local noise are regarded as the effective noise, which is assumed to be statically independent of the target signal.

(3) The target-signal components $s_i(t)$ observed at the outputs of the ith hydrophone is a linear time-invariant transformation of d(t), the target-signal component observed at the output of an ideal isotropic hydrophone located at the origin of the coordinates. The target direction is known, together with its autocorrelation function (but not necessarily its power level).

(4) The statistics of the noise field are completely unknown. Interferences may be present, but this is unknown. If they are present, their directions are unknown.

(5) The wavefronts of target and interferences are regarded as plane over the dimensions of the array.

(6) The processor is a directional array whose gain is maximized in the direction from which the target is expected to come.

In Chapter V, in order to analyze the performance of the proposed processor, it was further assumed that

(7) The array is linear and consists of equal spaced hydrophones.

(8) The ambient noise is statistically independent from hydrophone to hydrophone.

(9) The input processes are band limited and of similar spectra.

b) Summary of Results

A mathematical model has been developed to describe the characteristics of the input processes and the processing mechanisms. This model has been used to examine the array processor when the filter coefficients are adjusted iteratively so as to optimize the processor in accordance with the following performance measures:

(1) Minimum mean-squared error between the beamformer output and the target signal for the filtering problem.

(2) Maximum signal-to-noise ratio at the processor output for the detection problem.

For a general array configuration consisting of individual filter on each hydrophone output, a post-summation filter, a square-law device, and an averaging filter, the optimum individual filters can be constructed by tapped-delay lines with the weights set to some optimal values. Although these optimal values cannot be determined without complete knowledge about both the target and the noise, methods of stochastic approximation can be applied to adjust the weights iteratively. The only information required in using the adaptive algorithms is the correlation functions between the wavefront and the various delayed signals.

The proposed algorithms have been shown to converge in mean square and in

probability as long as the second order statistics of the input processes are bounded. Explicit expressions for the rate of convergence are derived in terms of input statistics, various system parameters and training environment. The mean-squared error is found to decrease approximately as the first power of the adaptation time. The rate of convergence is essentially indifferent to the number of weights to be adjusted as our algorithm allows simultaneous adjustments. The size of error, however, depends on the total number of taps and the starting point. Ranges of the weighting sequence are determined to maintain stability of the adaptive loop. It is also of interest to note that there is no signal suppression phenomenon in using our algorithm and that the final system performance is independent of the signal power level.

Several partially effective techniques have been proposed to adjust time-varying parameters. It is also found that the ordinary methods of stochastic approximation can still provide convergent algorithms if the rate of parameter variation is sufficiently slow. Qualitative discussions are provided.

The performances of the proposed adaptive receiver are evaluated and compared with those of the non-adaptive systems. The whole system starts to be a conventional detector and is gradually transformed into a spare-time filter optimum in a predetermined direction. This optimum filter is shown to reduce disturbances coming from other directions. When a signal appears in this particular direction, a maximum response will be produced. In actual operation, the average bearing response can be obtained from a plot of the averaged squared output versus the looking angle of the array. In most practical situations narrow peaks are considered to be targets.

7.2 Suggestions for Future Research

The following problem areas have been suggested by the research reported here:

a) Applications in Other Areas

Much work remains to be done in other areas of application. New areas of

application should be explored both from a theoretical standpoint and from a practical one. Two important areas are seismic arrays and satellite communications.

In processing seismic data the direction of the source is generally known because of the impulse nature of the initial signal. The direction and nature of seismic noise are not easily determined. The iterative procedure suggested in this research could be used to minimize the effect of such noises.

The suggested system presented here might be used to improve the signal-to-noise ratio for communication signals received from transmitters located on deep space probes. Presumably, the direction of the source is known (e.g., the location of a satellite), but the characteristics of the interfering noises are unknown. The improvement offered by the array-processing system presented here, as compared with conventional systems, might be appreciable.

Detailed analysis of the above two areas will be very useful and important in understanding the performances of these adaptive systems.

b) Nonstationary Problems

The applicability of adaptive techniques to statistically nonstationary processes presents some highly challenging mathematical and statistical problems, and perhaps is the one in which the strongest applications of adaptive techniques will be made. In this research some procedures have been proposed. But they are applicable only to special cases. A generalized formulation to handle this problem would be highly desirable.

c) Automatic Recognition of Bearing Response

In applying the proposed algorithm to actual sonar systems, an operator is needed to interpret the bearing response. One would like to ask whether or not an automatic response reader can be constructed by studying the characteristics of directivity patterns and by developing some recognition algorithms.

APPENDIX A

THE OPTIMUM DETECTOR FOR DETECTION OF A GAUSSIAN SIGNAL IN GAUSSIAN NOISE BACKGROUND

Suppose that the array consists of K hydrophones, and that the received signal at the i^{th} hydrophone is $x_i(t)$

$$x_i(t) = s_i(t) + n_i(t) \ , \quad i = 1,2,\cdots, K \tag{A-1}$$

where $s_i(t)$ is the signal that would be observed at the i^{th} hydrophone if there were no noise, and $n_i(t)$ is the noise which includes both ambient noise and interferences. Both $s_i(t)$ and $n_i(t)$ are assumed to be Gaussian random processes with zero mean and so is the input $x_i(t)$. If the spectrum of $x_i(t)$ is limited to frequencies below ω_o cps, and the $x(t)$ are observed over an interval T, such that $\omega_o T >> 1$, then $x_i(t)$ can be expanded in a Fourier series

$$x_i(t) = \sum_{n=-\omega_o T}^{\omega_o T} x_i(n) e^{j2\pi nt/T} \tag{A-2}$$

where $x_i(n)$ are complex Fourier coefficients satisfying $x_i(-n) = x_i^*(n)$ and where the asterisk stands for complex conjugate. It is seen that all the available information about the signals received by the entire array is contained in the set of vectors

$$\underline{X}(n) = \begin{matrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_K(n) \end{matrix} \tag{A-3}$$

Following [6] and [7], we assume that $\underline{X}(n)$ and $\underline{X}(m)$ are statistically independent for $n \neq m$. By the same token, we let the signal component of $x_i(t)$

be given by

$$s_i(t) = \sum_{n=-\omega_o T}^{\omega_o T} s_i(t) e^{j2\pi nt/T} \tag{A-4}$$

so that the signal at all hydrophones is represented by

$$\underline{S}(n) = \begin{bmatrix} s_1(n) \\ s_2(n) \\ \cdot \\ \cdot \\ \cdot \\ s_K(n) \end{bmatrix} \tag{A-5}$$

Here again we assume that $\underline{S}(n)$ is independent of $\underline{S}(m)$ for $n \neq m$.

The optimum detector is known to be the likelihood ratio detector, which determines the presence or absence of a target by comparing the likelihood ratio

$$LR = \frac{f_s(\underline{X})}{f_N(\underline{X})} \tag{A-6}$$

to a fixed threshold. Here $f_s(\underline{X})$ is the conditional probability density function of the received samples (over all hydrophones and over all frequencies) when signal is assumed to be present; similarly $f_N(\underline{X})$ is the conditional probability density function when signal is assumed to be absent. Since $\underline{X}(-n) = \underline{X}^*(n)$, and $\underline{X}(n)$ and $\underline{X}(m)$ are independent for $n \neq m$, Eq. (A-6) can be written as

$$LR = \prod_{n=1}^{\omega_o T} \frac{f_s[\underline{X}(n)]}{f_N[\underline{X}(n)]} \tag{A-7}$$

Now, define the signal and noise covariance matrices at each frequency by

$$\underline{P}(n) = \langle \underline{X}^*(n)\ \underline{X}^T(n) \rangle_s \tag{A-8}$$

$$\underline{Q}(n) = \langle \underline{X}^*(n)\ \underline{X}^T(n) \rangle_N \tag{A-9}$$

where the superscript T refers to transposition and the symbol $\langle\ \rangle_N$ means ensemble average subject to the noise-only hypothesis. Then the conditional probability density functions appearing in Eq. (A-7) can be expressed as

$$f_N\,[\underline{X}(n)] = C_N \exp\,[-\,\underline{X}^{*T}(n)\;\underline{Q}^{-1}(n)\;\underline{X}(n)] \tag{A-10}$$

$$f_s\,[\underline{X}(n)] = C_{N+s} \exp\,[-\,\underline{X}^{*T}(n)\;\{\underline{P}(n) + Q(n)\}^{-1}\;\underline{X}(n)] \tag{A-11}$$

so that the likelihood ratio is

$$LR = \prod_{n=1}^{\omega_o T} \frac{C_{N+s}(n)}{C_N(n)} \exp\,[\underline{X}^{*T}(n)\;\{\underline{Q}^{-1}(n) - [\underline{P}(n) + \underline{Q}(n)]^{-1}\}\underline{X}(n)] \tag{A-12}$$

where the c's are the normalizing constant of the Gaussian distribution. We further assume that the signal originates from a source sufficiently remote from the array so that the wavefront is plane as it approaches the receiver. Referring to Fig. 1, we have

$$\underline{P}(n) = \phi_d(n)\;\underline{a}(n)\;\underline{a}^{*T}(n) \tag{A-13}$$

where $\phi_d(n)$ is the signal spectral density at frequency n ,

$a_i(n) = \exp\,[\,j\,\frac{2\pi n \tau_i}{T}\,]$, and τ_i is the delay at the i^{th} hydrophone. Since $\underline{P}(n)$ is now of rank 1, the inversion of the second term in the brackets can be written [7]

$$\begin{aligned}[\underline{Q} + \underline{P}]^{-1} &= [Q + \phi a\; a^{*T}]^{-1} \\ &= Q^{-1} - \frac{\underline{Q}^{-1}\;\underline{P}\;\underline{Q}^{-1}}{1 + \phi\;\underline{a}^{*T}\;\underline{Q}^{-1}\;\underline{a}}\end{aligned} \tag{A-14}$$

Using Eq. (A-14), one finds that the logarithm of the likelihood ratio is given by

$$\log LR = C + \sum_{n=1}^{\omega_o T} \frac{\underline{X}^{*T}(n)\;\underline{Q}^{-1}(n)\;\underline{P}(n)\;\underline{Q}^{-1}(n)\;\underline{X}(n)}{1 + \phi_d(n)\;\underline{a}^{*T}(n)\;\underline{Q}^{-1}(n)\;\underline{a}(n)} \tag{A-15}$$

where

$$C = \log \prod_{n=1}^{\omega_o T} C_{s+N}(n)/C_N(n)$$

Since $\underline{Q}^T(n) = \underline{Q}^*(n)$, the quadratic form appearing in Eq. (A-15) can be written in the form

$$\underline{X}^{*T}\,\underline{Q}^{-1}\,\underline{P}\,\underline{Q}^{-1}\,\underline{X} = \underline{X}^{*T}\,\underline{Q}^{-1}\,\phi_d\,\underline{a}\,\underline{a}^{*T}\,\underline{Q}^{-1}\,\underline{X}$$

$$= \phi_d\,(\underline{X}^T\,\underline{Q}^{*-1}\,\underline{a}^*)^*\,(\underline{Q}^{*-1}\,\underline{a}^*)^T\,\underline{X}$$

$$= \phi_d\,(\underline{X}^T\,\underline{Q}^{*-1}\,\underline{a}^*)^*\,(\underline{X}^T\,\underline{Q}^{*-1}\underline{a}^*) = \phi_d\,|\underline{X}^T\,\underline{Q}^{*-1}\,\underline{a}^*|^2 \qquad \text{(A-16)}$$

and therefore

$$\log LR = C + \sum_{n=1}^{\omega_o T} |\underline{X}^T(n)\,\underline{H}(n)|^2\,G_L^2(n) \qquad \text{(A-17)}$$

where

$$\underline{H}_o(n) = \underline{Q}^{*-1}(n)\,\underline{a}^*(n) \qquad \text{(A-18)}$$

$$G_L^2(n) = \phi_d/[\,1 + \phi_d(n)\,\underline{a}^{*T}(n)\,\underline{Q}^{-1}(n)\,\underline{a}(n)] \qquad \text{(A-19)}$$

$\underline{H}_o(\omega)$ and $G_L(\omega)$ are the optimum individual and post-summation filters, respectively, as referred in Fig. 3.

If we let $n = \omega_n$, $\Delta\omega = \frac{1}{2\pi T}$, the summation appearing in Eq. (A.17) can be transformed into an integral for large T

$$\sum_{n=1}^{\omega_o T} |\underline{X}\,(\omega_n)^T\,\underline{H}(\omega)\,G_L|^2$$

$$\simeq \frac{1}{2\pi T}\int_{-\infty}^{\infty} |\underline{X}^T(\omega)\,\underline{H}(\omega)\,G_L(\omega)|^2\,d\omega$$

$$= \frac{1}{2\pi T}\int_{-\infty}^{\infty} |\sum_{\ell=1}^{K} G_L(\omega)\,H_\ell(\omega)\,X_\ell(\omega)|^2\,d\omega$$

$$= \frac{1}{T}\int_{o}^{T} |\sum_{\ell=1}^{K} \{F^{-1}[G_L(\omega)\,H_\ell(\omega)\,X_\ell(\omega)]\}|^2\,dt \qquad \text{(A-20)}$$

The last expression is obtained by invoking Parseval's theorem. Eq. (A.20) and hence Eq. (A.17) can be implemented in a form represented either by Fig. A-1a or by Fig. A-1b. These two structures are equivalent, but the latter is drawn here for future comparison. $\underline{\Phi}_{nn}(\omega)$ is just the noise matrix $\underline{Q}(\omega)$ and used here to make the nomenclature consistent with our previous developments.

We shall now consider other performance criteria for the array system. Referring to the general array configuration Fig. 3, and combining the post-summation and the individual filters to make $G(\omega) = 1$, we see from Sect. 5.2 that the signal-to-noise ratio at the detector output is

$$\mathrm{SNR} = \frac{\frac{1}{2}\ \frac{T}{\pi} \int_{-\infty}^{\infty} [\underline{H}^T\ \underline{\Phi}_{ss}\ \underline{H}^*] d\omega}{\left[\int_{-\infty}^{\infty} [H^T\ \underline{\Phi}_{xx}\ \underline{H}^*]^2 d\omega \right]^{1/2}} = \frac{N}{D} \tag{A-21}$$

Assume that $\underline{H}_M(\omega)$ maximizes (A.21), and let

$$\underline{H}(\omega) = \underline{H}_M(\omega) + \varepsilon \underline{h}(\omega) \tag{A-22}$$

where $\underline{h}(\omega)$ is an arbitrary vector function of ω . Eq. (A-21) must now have a maximum at $\varepsilon = 0$ if $\underline{H}_M$ is in fact optimum. That is,

$$\frac{d\mathrm{SNR}}{d\varepsilon}\bigg|_{\varepsilon=0} = 0 = \frac{D_M}{2} \quad \frac{T}{\pi} \int_{-\infty}^{\infty} [\underline{H}_M^T\ \underline{\Phi}_{ss}\ \underline{h}^* + \underline{h}^T\ \underline{\Phi}_{ss}\ H_M^*] d\omega$$

$$- N_M\ \frac{1}{2}\ D_M^{-1} \int_{-\infty}^{\infty} 2[\underline{H}_M^T\ \underline{\Phi}_{xx}\ \underline{H}_M^*]\ [\underline{h}^T\ \underline{\Phi}_{xx}\ \underline{H}_M^* + \underline{H}_M^T\ \underline{\Phi}_{xx}\ \underline{h}^*] d\omega$$

$$= \int_{-\infty}^{\infty} \underline{H}_M^T\ [c^2\ \underline{\Phi}_{ss} - \underline{\Phi}_{xx}\ \underline{H}_M^*\ \underline{H}_M^T\ \underline{\Phi}_{xx}] \underline{h}^*\ d\omega$$

$$+ \int_{-\infty}^{\infty} \underline{h}^T\ [c^2\ \underline{\Phi}_{ss} - \underline{\Phi}_{xx}\ \underline{H}_M^*\ \underline{H}_M^T\ \underline{\Phi}_{xx}]\ H_M^*\ d\omega \tag{A-23}$$

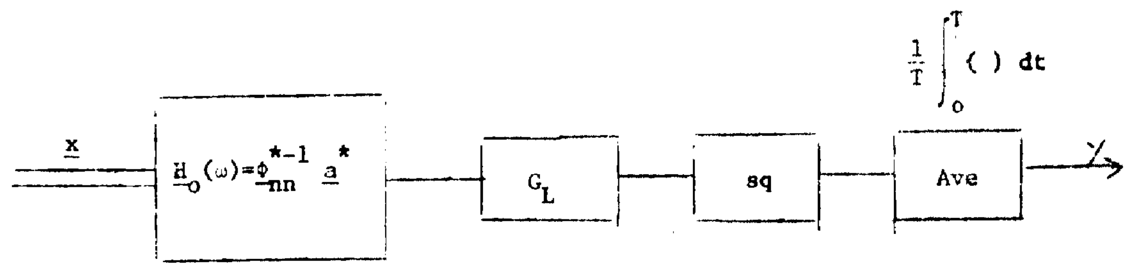


Fig. A - 1a

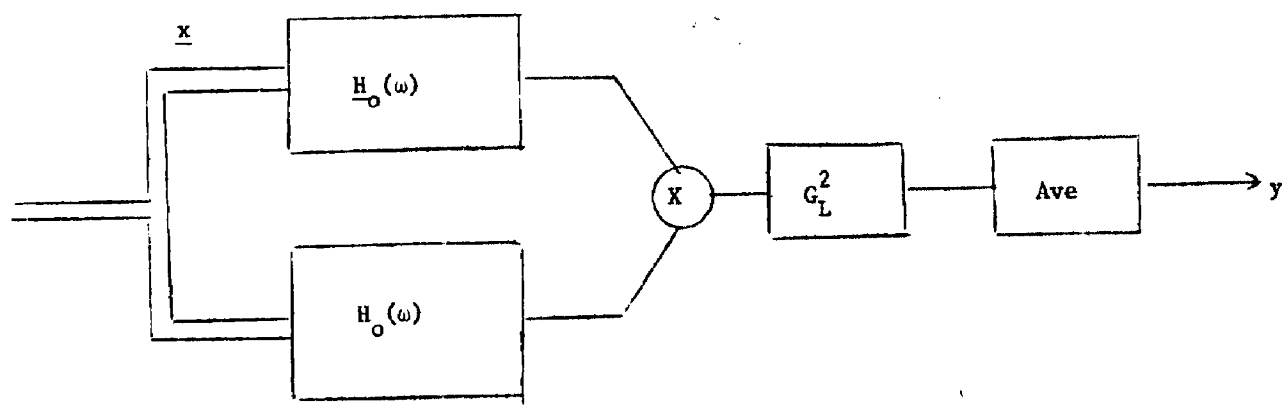


Fig. A - 1b.

Fig. A-1 Two Equivalent Structures of Likelihood Ratio Detector.

where D_M , N_M are the optimum value of denominator and numerator, respectively, of Eq. (A-21)

$$c^2 = \frac{D_M^2 \sqrt{T/\pi}}{2 N_M}$$

is actually arbitrary since any constant multiplier in Eq. (A-21) cannot effect the condition for a maximum. The two integrals of Eq. (A-23) are equivalent by virtue of the fact that the spectral matrix is Hermitian

$$\Phi^{*T} = \Phi \tag{A-24}$$

and [24]

$$\int_{-\infty}^{\infty} H_1^T \Phi H_2^* d\omega = \int_{-\infty}^{\infty} H_2^T \Phi H_1^* d\omega \tag{A-25}$$

Hence

$$\left.\frac{dSNR}{d\varepsilon}\right|_{\varepsilon=0} = \int_{-\infty}^{\infty} \underline{H}^T[c^2 \underline{\Phi}_{ss} - \underline{\Phi}_{xx} \underline{H}_M^* \underline{H}_M^T \underline{\Phi}_{xx}]h^* d\omega = 0 \tag{A-26}$$

Every component of $\underline{h}^*$ is arbitrary, however, and Eq. (A-26) will not be satisfied for all $\underline{h}^*$ unless

$$\underline{H}_M^T [c^2 \underline{\Phi}_{ss} - \underline{\Phi}_{xx} \underline{H}_M^* \underline{H}_M^T \underline{\Phi}_{ss}] = 0 \tag{A-27}$$

Taking c = 1 and using Eq. (2.1-2), i.e.,

$$\underline{\Phi}_{ss} = \phi_d \underline{a} \underline{a}^{*T}$$

Eq. (A-27) reduces to

$$\underline{H}_M = [\underline{\Phi}_{xx}]^{-1} \sqrt{\phi_d} \underline{a}^* = \phi_d^{-\frac{1}{2}} [\underline{\Phi}_{xx}^*]^{-1} \phi_d \underline{a}^* \tag{A-28}$$

But,

$$[\underline{\Phi}_{xx}]^{-1} \phi_d \underline{a}^* = [\underline{\Phi}_{nn}^* + \phi_d \underline{a}^* \underline{a}^T]^{-1} \phi_d \underline{a}^*$$

$$= [\underline{\Phi}_{nn}^{*-1} - \frac{\phi_d \underline{\Phi}_{nn}^{*-1} \underline{a}^* \underline{a}^T \underline{\Phi}_{nn}^{*-1}}{1 + \phi_d \underline{a}^T \underline{\Phi}_{nn}^{*-1} \underline{a}^*}] \phi_d \underline{a}^*$$

$$= \underline{\Phi}_{nn}^{*-1} \underline{a}^* \frac{\phi_d}{1 + \phi_d \underline{a}^T \underline{\Phi}_{nn}^{*-1} \underline{a}^*} \tag{A-29}$$

Eq. (A-28) can be rewritten as

$$\underline{H}_M = [\ \underline{\Phi}_{nn}^{*-1}\ ]\ \underline{a}^* \frac{\Phi_d^{\frac{1}{2}}}{1 + \Phi_d\ \underline{a}^T\ \underline{\Phi}_{nn}^{-1}\ \underline{a}^*} = \underline{H}_o(\omega)\ G_M(\omega) \tag{A-30}$$

where

$$G_M = \frac{\Phi_d^{\frac{1}{2}}}{1 + \Phi_d\ \underline{a}^T\ \underline{\Phi}_{nn}^{-1}\ \underline{a}^*} = \frac{G_L^2}{\Phi_d^{\frac{1}{2}}} \tag{A-31}$$

is the post-summation filter maximizing the signal-to-noise ratio at the detector output. If one wishes to minimize the mean squared error between the signal wavefront and the summer output

$$\overline{e^2} = \overline{[d(t) - z(t)]^2} = \overline{z^2} + \overline{d^2} - 2\ R_{dz}(0) \tag{A-32}$$

then, by using Eq. (5.2-9)

$$\overline{z^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_z(\omega)d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \underline{H}^T\ \underline{\Phi}_{xx}\ \underline{H}^*\ d\omega \tag{A-33}$$

$$R_{dz}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_d\ \underline{a}^{*T}\ \underline{H}^*\ d\omega \tag{A-34}$$

one obtains

$$\overline{e^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} [\underline{H}^T\ \underline{\Phi}_{xx}\ \underline{H}^* + \Phi_d(\omega) - 2\Phi_d(\omega)\underline{a}^{*T}\ \underline{H}^*]d\omega \tag{A-35}$$

Employing the same techniques of calculus of variation, Eq. (A-35) is minimized by choosing

$$\underline{H}_m = [\ \underline{\Phi}_{xx}\ ]^{-1}\ \underline{a}^*\ \Phi_d = \underline{\Phi}_{nn}^{*-1}\ \frac{\Phi_d}{1 + \Phi_d\ \underline{a}^{*T}\ \underline{\Phi}_{nn}^{-1}\ \underline{a}}$$

$$= \underline{H}_o\ G_m \tag{A-36}$$

where

$$G_m = \frac{\Phi_d}{1 + \Phi_d(\omega)\ \underline{a}^{*T}\ \underline{\Phi}_{nn}^{-1}\ \underline{a}} = \Phi_d^{\frac{1}{2}}\ G_M \tag{A-37}$$

is the post-summation filter minimizing the mean square error.

Thus, the optimum individual filter for our signal model is

$$\underline{H}_o(\omega) = \underline{\Phi}_{nn}^{*-1}(\omega)\ \underline{a}^*(\omega) \tag{A-38}$$

and is invariant under changes of optimization criteria. Only the optimum post-summation filter $G(\omega)$ needs to be modified according to

$$G_m = G_L^2 = G_M\, \phi_d^{\frac{1}{2}} \tag{A-39}$$

APPENDIX B

PROOF OF THEOREM 1

Theorem 1: Let $\gamma_1, \gamma_2, \ldots$ be a sequence of positive numbers such that

(A 1) $\lim_{j\to\infty} \gamma_j = 0$ (B-1a)

(A 2) $\sum_{j=1}^{\infty} \gamma_j = \infty$ (B-1b)

(A 3) $\sum_{j=1}^{\infty} \gamma_j^2 < \infty$ (B-1c)

(A 4) $\gamma_j > 0$ (B-1d)

Let the following conditions be satisfied

(B) $$\inf_{\varepsilon < \|\underline{c} - \underline{c}_{op}\| < \frac{1}{\varepsilon}} (\underline{c} - \underline{c}_{op})^T E(\nabla_c Q(\underline{x}|\underline{c})) > 0$$ (B-2)

$\varepsilon > 0$

(C) $$E(\nabla_c^T Q(\underline{x}|\underline{c}) \nabla_c Q(\underline{x}|\underline{c})) \le d(\underline{c}_{op}^T \underline{c}_{op} + \underline{c}^T \underline{c})$$ (B-3)

for all $\underline{c}$ in a bounded se[illegible]

and $d > 0$.

Then the seq[illegible]

$$\underline{c}_{j+1} = \underline{c}_j - [illegible]$$ (B-4)

converges with [illegible]

Proof: [illegible]

$$\underline{c}_{j+1} - \underline{c}_{op} = \underline{c}_j - [illegible]$$ (B-5)

where, for simplicity, [illegible]

Best Available Copy

Squaring Eq. (B-5)

$$(\underline{c}_{j+1} - \underline{c}_{op})^T (\underline{c}_{j+1} - \underline{c}_{op}) = (\underline{c}_j - \underline{c}_{op})^T (\underline{c}_j - \underline{c}_{op})$$

$$- 2\gamma_j (\underline{c}_j - \underline{c}_{op})^T \underline{\nabla} Q$$

$$+ \gamma_j^2 \underline{\nabla}^T Q \underline{\nabla} Q$$

and taking the conditional mathematical expectation for given $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j$, we obtain

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2 \,|\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j\}$$

$$= ||\underline{c}_j - \underline{c}_{op}||^2 - 2\gamma_j (\underline{c}_j - \underline{c}_{op})^T E\{\underline{\nabla} Q\}$$

$$+ \gamma_j^2 E\{\underline{\nabla}^T Q \underline{\nabla} Q\} \tag{B-6}$$

From condition (C), Eq. (B-6) becomes

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2|\underline{c}_1, \ldots \underline{c}_j\} \leq ||\underline{c}_j - \underline{c}_{op}||^2 - 2\gamma_j E(\underline{c}_j - \underline{c}_{op})^T \nabla Q\}$$

$$+ \gamma_j^2 d (\underline{c}_{op}^T \underline{c}_{op} + \underline{c}_j^T \underline{c}_j) \tag{B-7}$$

Using condition (B), Eq. (B-7) is reduced to

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2|\underline{c}_1, \ldots, \underline{c}_j\}$$

$$< ||\underline{c}_j - \underline{c}_{op}||^2 (1 + \gamma_j^2 d) + 2\gamma^2{}_j d \underline{c}^T \underline{c}_{op} \tag{B-8}$$

Let $\underline{Z}_j = ||\underline{c}_j - \underline{c}_{op}||^2 \prod_{k=j}^{\infty} (1 + \gamma_k^2 d)$

$$+ \sum_{k=j}^{\infty} 2d\gamma_k^2 \underline{c}^T \underline{c}_{op} \sum_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \tag{B-9}$$

Then $\underline{Z}_{j+1} = ||\underline{c}_{j+1} - \underline{c}_{op}||^2 \prod_{k=j+1}^{\infty} (1 + \gamma_k^2 d)$

$$+ \sum_{k=j+1}^{\infty} 2d\gamma_k^2 \underline{c}^T \underline{c}_{op} \prod_{m=k+1}^{\infty} (1 + \gamma_m^2 d) \tag{B-10}$$

Taking the conditional mathematical expection for given $\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_j$, we have

$$E\{Z_{j+1}|\underline{c}_1,\ldots,\underline{c}_j\} = E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2|\underline{c}_1,\ldots\underline{c}_j\}\prod_{k=j+1}^{\infty}(1 + \gamma_k^2\, d)$$

$$+ \sum_{k=j+1}^{\infty} 2d\gamma_k^2\, \underline{c}^T\, \underline{c}_{op} \prod_{m=k+1}^{\infty} (1 + \gamma_m^2\, d)$$

$$\leq\ ||\underline{c}_j - \underline{c}_{op}||^2\ (1 + d\gamma_j^2) + 2\gamma_j^2\ d\underline{c}^T\ \underline{c}_{op} \prod_{k=j+1}^{\infty}(1 + \gamma_k^2\, d)$$

$$+ \sum_{k=j+1}^{\infty} 2d\gamma_k^2\, \underline{c}^T\, \underline{c}_{op} \prod_{m=k+1}^{\infty} (1 + \gamma_m^2\, d)$$

$$= \underline{Z}_j$$

or $E\{\underline{Z}_{j+1}|\underline{c}_1,\ldots,\underline{c}_j\} \leq \underline{Z}_j$ (B-11)

Next, taking the conditional mathematical expectation for given $\underline{Z}_1,\ldots,\underline{Z}_j$ on both sides of Eq. (B-11), we have

$$E\{\underline{Z}_{j+1}|\underline{Z}_1,\ldots,\underline{Z}_j\} \leq \underline{Z}_j \qquad \text{(B-11a)}$$

Since $\underline{Z}_j = f\ (\underline{c}_1, \underline{c}_2,\ldots,\underline{c}_j)$.

Inequality (B-11a) shows that $\underline{Z}_j$ is a semimartingale, where

$$E\ \underline{Z}_{j+1} \leq E\ \underline{Z}_j \leq \ldots \leq E\ \underline{Z}_1 < \infty \qquad \text{(B-12)}$$

so that, according to the theory of semimartingales[1] the sequence $\underline{Z}_1$ converges with probability one, and hence by virtue of Eqs. (B-1b) and (B-1c) the sequence $(\underline{c}_j - \underline{c}_{op})$ also converges with probability one to some random number $\underline{\xi}$. It remains to show that $P(\xi = 0) = 1$. It is seen that from Eqs. (B-12), (B-9) and (B-1c) the sequence $E(\underline{c}_j - \underline{c}_{op})$ is bounded. Now taking the mathematical expectation on both sides of the inequality (B-7)

---

1 Doob, J. L., Stochastic Processes, John Wiley and Sons, N. Y., 1953

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2\} \le E\{||\underline{c}_j - \underline{c}_{op}||^2\} - 2\gamma_j\ E\{(\underline{c}_j - \underline{c}_{op})^T\ \nabla\ Q\}$$

$$+ \gamma^2{}_j\ d[\underline{c}_{op}{}^T\ \underline{c}_{op} + E(\underline{c}_j{}^T\ \underline{c}_j)]$$

and adding the first j inequalities together, we have by deduction

$$E\{||\underline{c}_{j+1} - \underline{c}_{op}||^2\} \le E\{||\underline{c}_1 - \underline{c}_{op}||^2 + \sum_{k=1}^{j} (d\underline{c}_{op}{}^T\ \underline{c}_{op}\gamma_k^2 + d\gamma_k^2\ E(\underline{c}_k^T\ \underline{c}_k))\}$$

$$- \sum_{k=1}^{j} 2\gamma_k\ E\{(\underline{c}_j - \underline{c}_{op})^T\ \nabla\ Q\} \qquad \text{(B-13)}$$

Since $E\{||\underline{c}_j - \underline{c}_{op}||^2\}$ is bounded and condition (B-1c) is fulfilled, from Eq. (B-13) it follows that

$$\sum_{k=1}^{\infty} \gamma_k\ E\{(c_j - c_{op})^T\ \nabla\ Q\} < \infty \qquad \text{(B-14)}$$

Using condition (B-1b), i.e., $\sum_{j=1}^{\infty} \gamma_j = \infty$ and noting (B-2)

$$\inf_{\varepsilon < ||\underline{c} - \underline{c}_{op}|| < \frac{1}{\varepsilon}} E\{(\underline{c} - \underline{c}_{op})^T\ \nabla\ Q\} > 0$$

We deduce from Eq. (B-14) that

$$E\{(\underline{c}_j - \underline{c}_{op})^T\ \nabla\ Q\} \to 0 \text{ with probability one for some sequence } N\ . \qquad \text{(B-15)}$$

Now, taking $E\{||\underline{c}_j - \underline{c}_{op}||^2\} \to \underline{\xi}$ with probability one, and comparing Eq. (B-15) with Eq. (B-2) we obtain

$$\underline{\xi} = 0 \text{ with probability one.} \qquad \text{(B-16)}$$

Therefore, algorithm (B-4) converges with probability one

$$P\ \lim_{j\to\infty} (\underline{c}_j - \underline{c}_{op}) = 0 \quad = 1 \qquad \text{(B-17)}$$

as well as in mean square sense, i.e.,

$$\lim_{j\to\infty} E\{||\underline{c}_j - \underline{c}_{op}||^2\} = 0 \qquad \text{(B-18)}$$

APPENDIX C

Some properties of Gamma functions

Since
$$\Gamma(\alpha + n) = (\alpha + n - 1)\,\Gamma(\alpha + n - 1)$$
$$= (\alpha + n - 1)(\alpha + n - 2)\,\Gamma(\alpha + n - 2)$$
$$= - - - -$$
$$= (\alpha + n - 1)(\alpha + n - 2) --- \alpha\Gamma(\alpha)$$

We have

$$\prod_{k=1}^{n}(\alpha + k - 1) = \alpha(\alpha + 1) --- (\alpha + n - 1)$$
$$= \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)} \tag{C-1}$$

Thus

$$\prod_{k=1}^{j}\left(1 - \frac{\lambda}{j+1}\right) = \frac{\prod_{k=1}^{j}(j+1 - \lambda)}{(j+1)!} = \frac{\Gamma(j+2 - \lambda)}{(j+1)!\;\Gamma(2 - \lambda)} \tag{C-2}$$

Eq. (C-2) can be approximated by using the formula*

$$\Gamma(x) = e^{-x}\,x^{x-\frac{1}{2}}\,(2\pi)^{\frac{1}{2}}\left\{1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320\,x^4} + 0\left(\frac{1}{x^5}\right)\right\}$$
$$\approx e^{-x}\,x^{x-\frac{1}{2}}\,(2\pi)^{\frac{1}{2}} \quad \text{for} \quad x \gg 1 \tag{C-3}$$

From Eq. (C-3) we can write for $j \gg 1$,

$$\Gamma(j + 2 - \alpha) \simeq e^{-(j+2-\alpha)}\,(j+2-\alpha)^{j+2-\alpha-\frac{1}{2}}\,(2\pi)^{\frac{1}{2}}$$
$$= e^{-(j+2-\alpha}\,(j+2-\alpha)^{j+\frac{3}{2}}\,(j+2-\alpha)^{-\alpha}\,(2\pi)^{\frac{1}{2}} \tag{C-4}$$

$$(j+1)! = \Gamma(j+2) \simeq e^{-(j+2)}\,(j+2)^{j+\frac{3}{2}}\,(2\pi)^{\frac{1}{2}} \tag{C-5}$$

Since

$$j + 2 - \alpha \simeq j + 2 \quad \text{if} \quad j \gg \alpha$$

---

* Whittaker and Waston, Modern Analysis, p. 253

we obtain from Eqs. (C-4) and (C-5)

$$\frac{\Gamma(j+2-\alpha)}{(j+1)!} = \frac{\Gamma(j+2-\alpha)}{\Gamma(j+2)} \approx (j+2-\alpha)^{-\alpha}$$

$$\approx \frac{1}{(j+1)^{\alpha}} \quad \text{if} \quad j >> 1 \quad \text{and} \quad j >> \alpha \tag{C-6}$$

Therefore, combining Eqs. (C-2) and (C-6) gives

$$\prod_{k=1}^{j} (1 - \frac{\lambda}{j+1}) \approx \frac{1}{\Gamma(2-\lambda)(j+1)^{\lambda}} \tag{C-7}$$

and furthermore,

$$\prod_{j=m}^{n}(1 - \frac{A}{j+1}) \approx \frac{m^{A}}{(n+1)^{A}} \tag{C-8}$$

APPENDIX D

EFFECT OF UNCERTAIN SIGNAL POWER ON THE FINAL VALUES OF THE GAINS

To illustrate the essential steps involved in studying the convergence properties of algorithm (3.5-1), we shall consider only the single gain case. The corresponding extensions to multiple gain case is straightforward but laborious due to matrix manipulations. Examples have been presented in Sect. 3.5.

Let the assumed signal correlation function $\hat{R}_s$ and the actual signal correlation function R_s be related by a multiplicative constant G_s

$$\hat{R}_s = G_s R_s \tag{D-1}$$

The single gain version of (3.5-1) is

$$\begin{aligned} c_{j+1} &= c_j + 2\gamma_j \hat{R}_{ds} - 2\gamma_j z_j n_j \\ &= c_j + 2\gamma_j G_s R_s - 2\gamma_j c_j x_j^2 \end{aligned} \tag{D-2}$$

since $z(t) = cn(t) = cx(t)$ in this simple case.

The optimum gain is known to be

$$\theta = c_{op} = (R_s + R_n)^{-1} R_s = \overline{x^2}^{\,-1} R_s \tag{D-3}$$

The average of Eq. (D-2) is then

$$\begin{aligned} \bar{c}_{j+1} &= (1 - 2\gamma_j \overline{x^2}) \bar{c}_j + 2\gamma_j G_s \overline{x^2}\, \theta \\ &= c_1 \prod_{k=1}^{j} (1 - 2\gamma_k \overline{x^2}) \\ &+ G_s \theta \sum_{k=1}^{j} 2\gamma_k \overline{x^2} \prod_{\ell=k+1}^{j} (1 - 2\gamma_\ell \overline{x^2}) \end{aligned} \tag{D-4}$$

If we set

$$\gamma_j = \frac{1}{2(j+1)\overline{x^2}} \tag{D-5}$$

Eq. (D-4) becomes

$$\overline{c}_{j+1} = \frac{c_1}{j+1} + G_s\theta \frac{1}{j+1} \tag{D-6}$$

which shows that the mean of c_j converges to $G_s\theta$ at the rate of j^{-1}

$$(\overline{c}_{j+1} - G_s\theta) = \frac{(c_1 - G_s\theta)}{j+1} \rightarrow 0 \quad \text{as} \quad j\rightarrow\infty \tag{D-7}$$

We shall consider $\overline{c^2_{j+1}}$.
Squaring (D-2)

$$c^2_{j+1} = (1 - 2\gamma_j\, x^2_j)^2\, c^2_j + 4\gamma_j{}^2\, G^2_s\, R^2_s$$

$$+ 4\gamma_j(1 - 2\gamma_j\, x^2_j)\, c_j\, G_s\, R_s$$

and taking the average yield

$$\overline{c^2_{j+1}} = (1 - 4\gamma_j\, \overline{x^2} + 4\gamma_j{}^2\, \overline{x^4})\, \overline{c_j{}^2}$$

$$+ 4\gamma_j\, G^2_s\, R^2_s + 4\gamma_j\, (1 - 2\gamma_j\, \overline{x^2})\, G_s\, R_s\, \overline{c}_j \tag{D-8}$$

If we let

$$4\gamma_j\, \overline{x^2} - 4\gamma_j{}^2\, \overline{x^4} = v_j \tag{D-9}$$

$$4\gamma_j{}^2\, G^2_s\, R^2_s + 4\gamma_j(1 - 2\gamma_j\, \overline{x^2})\, G_s\, R_s\, \overline{c}_j = u_j \tag{D-10}$$

Eq. (D-8) reduces to a simpler form

$$\overline{c^2_{j+1}} = (1 - v_j)\, \overline{c^2_j} + u_j$$

$$= c^2_1 \prod_{k=1}^{j} (1 - v_j) + \sum_{k=1}^{j} u_j \prod_{\ell=k+1}^{j} (1 - v_j) \tag{D-11}$$

From Eqs. (D-5) and (D-9) we can write

$$(1 - v_k) = 1 - (4\gamma_k\, \overline{x}^2 - 4\gamma_k{}^2\, \overline{x}^4) = (1 - \frac{A_1}{j+1})(1 - \frac{A_2}{j+1}) \tag{D-12}$$

where

$$A_1 = 1 - \sqrt{1 - a} \tag{D-13}$$

$$A_2 = 1 + \sqrt{1 - a} \tag{D-14}$$

$$a = \overline{x^4}/\overline{x^2} \tag{D-15}$$

Two approximations for the Gamma function will be used

$$\prod_{k=1}^{j} \left(1 - \frac{A}{k+1}\right) = \frac{\Gamma(j+1 - A)}{(j+1)!\,\Gamma(2-A)} \approx \frac{1}{\Gamma(2-A)(j+1)^A} \tag{D-16}$$

$$\prod_{k=m}^{n} \left(1 - \frac{A_1}{k+1}\right)\left(1 - \frac{A_2}{k+1}\right) = \frac{m^2}{(n+1)^2} \tag{D-17}$$

Using Eqs. (D-12) to (D-17), Eq. (D-11) becomes

$$\overline{c^2_{j+1}} \approx c_1^{\,2} \frac{1}{\Gamma(2-A_1)\Gamma(2-A_2)(j+1)^2} + \sum_{k=1}^{j} u_k \frac{(k+1)^2}{(j+1)^2} \tag{D-18}$$

But from Eqs. (D-10) and (D-5)

$$u_k = 4\gamma_k^{\,2}\, G_s^{\,2}\, R_s^{\,2} + 4\gamma_k(1 - 2\gamma_k\, \overline{x^2})\, G_s\, R_s\, \overline{c}_k$$

$$= \frac{G_s^{\,2}\, R_s^{\,2}}{(k+1)^2\, \overline{x^2}} + 4 \frac{1}{2(k+1)\, \overline{x^2}} \frac{k}{(k+1)} G_s\, R_s\, \overline{c}_k$$

$$= \frac{1}{(k+1)^2} \left[\frac{G_s^{\,2}\, R_s^{\,2}}{(\overline{x^2})^2} + \frac{2}{\overline{x^2}} k\, G_s\, R_s\, \overline{c}_k \right] \tag{D-19}$$

Substituting Eqs. (D-3) and (D-6) into Eq. (D-19) gives

$$u_k = \frac{1}{(k+1)^2} \left[G_s^{\,2}\, \sigma^2 + 2\, G_s^{\,2}\, \theta^2 \frac{k^2}{k+1} + 2G_s\theta\, c_1 \frac{k}{k+1} \right] \tag{D-20}$$

and Eq. (D-18) becomes

$$\overline{c_{j+1}^2} \approx \frac{c_1^{\ 2}}{\Gamma(2-A_1)\Gamma(2-A_2)(j+1)^2}$$

$$+ \frac{j}{(j+1)^2} G_s^{\ 2} \theta^2 + \frac{2G_s^{\ 2} \theta^2}{(j+1)^2} \sum_{k=1}^{j} \frac{k^2}{k+1}$$

$$+ \frac{2G_s \theta\ c_1}{(j+1)^2} \sum_{k=1}^{j} \frac{k}{k+1} \qquad \text{(D-21)}$$

For large j , we may write

$$\sum_{k=1}^{j} \frac{k^2}{k+1} \simeq \sum_{k=1}^{j} k = \frac{j(j+1)}{2} \qquad \text{(D-22)}$$

$$\sum_{k=1}^{j} \frac{k}{k+1} = \sum_{k=1}^{j} 1 = j \qquad \text{(D-23)}$$

Eq. (D-21) then becomes

$$c_{j+1}^2 \cong \frac{G_1^{\ 2}}{\Gamma(2-A_1)\Gamma(2-A_2)(j+1)^2}$$

$$+ \frac{j}{(j+1)^2} G_s^{\ 2} \theta^2 + G_s^{\ 2} \theta^2 + 2G_s\theta\ c_1 \frac{j}{(j+1)^2} \qquad \text{(D-24)}$$

The error variance in the parameter space is

$$\overline{(c_{j+1} - \theta)^2} = \overline{c_{j+1}^2} - 2\theta\ \bar{c}_{j+1} - \theta^2 \qquad \text{(D-25)}$$

which by utilizing Eqs. (D-24) and (D-6) reduces to

$$\overline{c_{j+1} - \theta)^2} \cong \frac{c_1^{\ 2}}{\Gamma(2-A_1)\Gamma(2-A_2)(j+1)^2}$$

$$+ \theta^2 \frac{1}{(j+1)^2} [j^2 (G_s - 1)^2 + j\ (3G_s^{\ 2} - 2G + 2) + G_s^{\ 2} + 1)]$$

$$+ 2c_1\ \theta \frac{1}{(j+1)^2} [j\ (G_s - 1) -1] \qquad \text{(D-26)}$$

or

$$\overline{(c_{j+1} - G_s\theta)^2} = \overline{c_{j+1}^2} - 2G_s\theta\,\overline{c}_{j+1} + G_s^2\,\theta^2$$

$$= \frac{c_1^2}{\Gamma(2-A_1)\Gamma(2-A_2)(j+1)^2}$$

$$+ G_s^2\,\theta^2\left[\frac{1}{(j+1)^2} + 2 - 2\,\frac{1}{j+1}\right]$$

$$+ 2G_s\theta\, c_1\left[\frac{1}{(j+1)^2} - \frac{1}{j+1}\right] \qquad \text{(D-27)}$$

and we have the asymptotical expression

$$\lim_{j\to\infty} \overline{(c_{j+1} - \theta)^2} = (G_s - 1)^2\,\theta^2 \qquad \text{(D-28)}$$

or

$$\lim_{j\to\infty} \overline{(c_{j+1} - G_s\theta)^2} = 0 \qquad \text{(D-29)}$$

Therefore, one can conclude that if the assumed signal power differs from the actual power by a multiplicative constant G_s, the gains adjusted according to algorithm (3.5-1) will converge in mean as well as in mean square to their optimum values multiplied by the same constant G_s.

APPENDIX E

GENERAL DYNAMIC METHODS OF STOCHASTIC APPROXIMATION

Recently Dupac [36] has proposed a dynamic stochastic approximation method for the particular case where the optimum of a production process moves linearly during the optimization period. This method essentially consists of a two-step, approximation procedure to be performed at each stage of the optimization process. The first step is designed to correct the time-varying trend of the parameters being estimated; the second step is made by means of an ordinary stochastic approximation procedure, based on the observation of a new sample. In [36], the parameters were assumed to vary linearly and the convergent conditions remain essentially unchanged from those of the stationary case.

Here we generalize the above method to include any nonlinear and coupling variations. Convergent conditions are modified accordingly and the proposed method is shown to reduce to Dupac's scheme as a special case.

a) Dupac's Method

Consider the problem of findi·g the extrema of functions of several variables

$$I = \overline{Q}\,(\underline{x}|\underline{c}) \tag{E-1}$$

where $\underline{x} = \{x_1, x_2, \ldots, x_n\}$ is a vector of random processes with distribution $P(\underline{x})$ and $\underline{c} = \{c_1, c_2, \ldots, c_n\}$ is a vector of parameters to be adjusted. When $P(\underline{x})$ is unknown, an algorithm derived from the method of stochastic approximation to obtain the set $\underline{c}_{op} = \underline{\theta}$ is

$$\underline{c}_{j+1} = \underline{c}_j - \gamma_j \nabla Q(\underline{x}_j|\underline{c}_j) \tag{E-2}$$

whose properties have been derived in Section 3.2 and Appendix B. It is assumed that $\underline{\theta}$ is time-invariant, i.e., $\underline{\theta}_j = \underline{\theta}$ for all j .

When the random environment is non-stationary, the optimum set $\underline{\theta}$

becomes a function of time index j. Its value at time j will be denoted by $\underline{\theta}_j$. Dupac [36] has considered the case where $\underline{\theta}_j$ is linearly (in his sense) time-varying.

$$\underline{\theta}_{j+1} - \underline{\theta}_j \left(1 + \frac{1}{j}\right) = 0 \left(\frac{1}{j^{\omega}}\right) \tag{E-3}$$

where $\omega > \alpha$, α being related to γ_j by $\gamma_j = 0 \left(\frac{1}{j^{\alpha}}\right)$.

Dupac's method is to estimate the unknown parameters $\underline{c}$ by

$$\underline{c}_{j+1} = \underline{c}_j \left(1 + \frac{1}{j}\right) - \gamma_j \nabla_c Q(\underline{x}_j | \underline{c}_j) \tag{E-4}$$

Algorithm (E-4) can be shown [36] to converge with probability one

$$P \left\{ \lim_{j \to \infty} (\underline{c}_j - \underline{\theta}_j) = 0 \right\} = 1 \tag{E-5}$$

as well as in mean square

$$\lim_{j \to \infty} E \left\{ ||\underline{c}_j - \underline{\theta}_j||^2 \right\} = 0 \tag{E-6}$$

under the following conditions

(A) $\gamma_j = \frac{\gamma}{j^{\alpha}}$, $\frac{1}{2} < \alpha \leq 1$, $\gamma > 0$ (E-7)

(B) There exist constants K_{ℓ} and K_u,

$0 < K_{\ell} < K_u < \infty$, such that

$$K_{\ell} ||\underline{c}_j - \underline{\theta}_j||^2 \leq (\underline{c}_j - \underline{\theta}_j)^T \overline{\nabla_c Q(\underline{x}_j | \underline{c}_j)} \leq K_u ||\underline{c}_j - \underline{\theta}_j||^2 \tag{E-9}$$

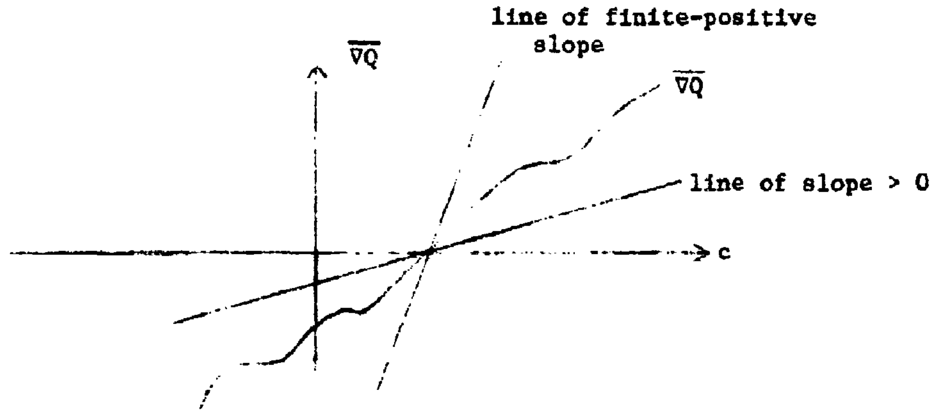
(C) For all values of $\underline{c}$,

$$\mathrm{Var}\,[\nabla^T Q\, \nabla Q] \leq \sigma^2 < \infty \tag{E-9}$$

where $\nabla Q \triangleq \nabla_c Q(\underline{x}|\underline{c})$ for simplicity.

Condition (B) and (C) are equivalent to conditions (3.2-14) and (3.2-15). Condition (B) simply says that $\overline{\nabla Q}$ must lie between two planes, one of positive slope and the other of finite-positive slope. This condition in one-

dimensional case is illustrated below



Condition (C) says that the variance of $\{\nabla^T Q \nabla Q\}$ is finite while Eq. (3.2-15) says that its expected value rather than its variance is finite. The conditions imposed here on the behavior of ∇Q are somewhat stronger than those for the ordinary methods of stochastic approximation.

b) <u>The General Method</u>

In this section we shall relax the restriction that the parameters vary linearly during the adaptation period. It is assumed that the law governing the variation is known, although the sequence to be estimated is unknown.

<u>Theorem 3</u>: Let the variation of $\underline{\theta}$ be governed by a known operator such that

$$\underline{\theta}_{j+1} = L\,(\underline{\theta}_j, j) \tag{E-10}$$

Define the following quantities

$$\underline{G}_j = \underline{G}(\underline{\theta}_j, j) = (\text{grad } L_j^T)^T \tag{E-11}$$

$$\lambda_j = \sup_{\text{all } \underline{\theta}} \{\text{eigenvalues of } \underline{G}_j\} \tag{E-12}$$

$$I_j = \overline{Q}_j = E\{Q(\underline{x}_j | c_j)\} \tag{E-13}$$

$$E\{\nabla_c Q_j | \underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{j-1}\} = \underline{m}_j \tag{E-14}$$

$$E\{\nabla_c{}^T Q_j \; \nabla_c Q_j | \underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{j-1}\} \le \underline{m}_j^T \underline{m}_j + d_j^2 \sigma^2 \tag{E-15}$$

$$\gamma_j = \frac{\gamma}{j^\alpha} \tag{E-16}$$

$$\lambda_j = \frac{\lambda}{j^\beta} \tag{E-17}$$

$$d_j = \frac{d}{j^\delta} \tag{E-18}$$

Then the algorithm

$$\underline{c}_{j+1} = L(\underline{c}_{j,j}) - \gamma_j \nabla_c Q(\underline{x}_j | \underline{c}_j) \tag{E-19}$$

converges in the sense of Eqs. (E-5) to (E-6) under the following conditions

A. $\frac{1}{2} - \delta < \alpha \le 1 + \beta$

$0 < \beta$

$$0 < \delta < \frac{1}{2} \tag{E-20}$$

B. There exist constants K_ℓ and K_u , $0 < K_\ell < K_u < \infty$, such that

$$K_\ell ||\underline{c}_j - \underline{\theta}_j||^2 \le (\underline{c}_j - \underline{\theta}_j)^T \; \overline{\nabla_c Q_j} \; \le K_u ||\underline{c}_j - \underline{\theta}_j||^2 \tag{E-21}$$

for $j=1, 2, \ldots$

C. For all values of $\underline{c}$

$$\text{Var}\,[\nabla_c{}^T Q \; \nabla_c Q] \le c^2 < \infty \tag{E-22}$$

Note that conditions (E-18) and (E-19) are just conditions (E-8) and (E-9), but Eq. (E-7) is replaced by Eq. (E-20) to take account of the time-varying effect.

<u>Proof</u>:

From Eqs. (E-10) and (E-19) the estimation error equation can be written in the form

$$\underline{e}_{j+1} = \underline{c}_{j+1} - \underline{\theta}_{j+1} = L(\underline{c}_{j,j}) - L(\underline{\theta}_{j,j}) - \gamma_j \nabla_c Q_j$$

$$= \underline{G}_j \; \underline{e}_j - \gamma_j \nabla_c Q_j \tag{E-23}$$

where $\underline{G}_j$, defined by Eq. (E-11), is a nonsingular matrix.

Take the inner product of Eq. (E-23)

$$\underline{e}_{j+1}^T \; \underline{e}_{j+1} = \underline{e}_j^T \, \underline{G}_j^T \, \underline{G}_j \, \underline{e}_j - 2\gamma_j \, \underline{e}_j^T \, \underline{G}_j^T \, \nabla_c Q_j + \gamma_j^2 \, Q_j \nabla_c Q_j \tag{E-24}$$

Note that $\underline{e}_j$ is a function of $\underline{x}_j$, $j = 1,2,\ldots$, and hence a random variable. Taking the conditional mathematical expectations of Eq. (E-24) and using the definitions Eqs. (E-12) to (E-15), we obtain

$$\begin{aligned} & E\{\underline{e}_{j+1}^T \, \underline{e}_{j+1} | \underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{j-1}\} \\ & \leq \lambda_j^2 \, E\{\underline{e}_j^T \, \underline{e}_j | \underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{j-1}\} \\ & \quad - 2\gamma_j \, \lambda_j \, \underline{m}_j^T \, \underline{e}_j + \gamma_j^2 \, [\underline{m}_j^T \, \underline{m}_j + d_j^2 \, \sigma^2] \end{aligned} \tag{E-25}$$

From Eq. (E-21) we can write

$$\underline{m}_j^T \, \underline{e}_j \geq K_\ell ||\underline{e}_j||^2 \tag{E-26}$$

$$\underline{m}_j^T \, \underline{m}_j \leq K_u ||\underline{e}_j||^2 \tag{E-27}$$

and thus Eq. (E-25) becomes

$$\begin{aligned} V_{j+1} & \leq \lambda_j^2 \, V_j - 2\gamma_j \, \lambda_j \, K_\ell V_j + \gamma_j^2 \, K_u V_j + \gamma_j^2 \, d_j^2 \, \sigma^2 \\ & = \lambda_j^2 \, (1 - 2\gamma_j \, k_j + \gamma_j^2 \, K_j) V_j + \gamma_j^2 \, d_j^2 \, \sigma^2 \end{aligned} \tag{E-28}$$

where

$$V_j = E\{\underline{e}_j^T \, \underline{e}_j | \underline{c}_1, c_2, \ldots, \underline{c}_{j-1}\} \tag{E-29}$$

$$k_j = K_\ell \lambda_j^{-1} \tag{E-30}$$

$$K_j = K_u \lambda_j^{-2} \tag{E-31}$$

Since the sequence γ_j is monotonic decreasing, there is J such that

$$(1 - 2\gamma_j k_j + \gamma_j^2 K_j) \leq [1-(2-\varepsilon_j) k_j \gamma_j] \tag{E-32}$$

and

$$[1 - (2 - \varepsilon_j)k_j\gamma_j] > 0 \tag{E-33}$$

$$0 < \varepsilon_j < 2 \quad \text{for} \quad j \geq J \tag{E-34}$$

Thus, for $j > J$, inequality Eq. (E-28) may be weakened to

$$V_{j+1} \leq \lambda_j^2[1-(2 - \varepsilon_j)k_j\gamma_j] + \gamma_j^2\, d_j^2\, \sigma^2 \tag{E-35}$$

The term in the bracket is positive for $j > J$, and we may start with $j = J$ and use this recursive inequality repeatedly to obtain

$$V_j \leq V_J B_{J-1,\, j-1} \prod_{\ell=J}^{j-1} \lambda_\ell^2 + \sigma^2 \sum_{\ell=J}^{j-1} \gamma_\ell^2\, d_\ell^2\, B_{\ell,J-1} \prod_{m=\ell}^{j-1} \lambda_m^2 \tag{E-36}$$

in which

$$B_{i,j} = \prod_{\ell=i+1}^{j} (1 - \gamma_\ell\, \xi_\ell)\ , \quad 0 \leq i \leq j \tag{E-37}$$
$$= 0 \qquad\qquad\qquad , \quad i > j$$

and

$$\xi_j = k_j(2 - \varepsilon_j) > 0 \tag{E-38}$$

By taking the logarithm of both sides of Eq. (E-37) and using the inequality

$$\log\, (1 - \gamma_\ell\xi_\ell) \leq -\gamma_\ell\xi_\ell \tag{E-39}$$

one may show that

$$B_{i,j} \leq \exp \left\{ - \sum_{\ell=i+1}^{j} \gamma_\ell\xi_\ell \right\}, \tag{E-40}$$

Therefore, it is necessary to have

$$\sum_{j=1}^{\infty} \gamma_j\xi_j = \infty \tag{E-41}$$

and

$$\sum_{j=1}^{\infty} \lambda_j^2 < \infty \tag{E-42}$$

to make the first term on the right-hand side of Eq. (E-36) vanish

$$\lim_{j\to\infty} B_{J-1,j-1} = 0 \tag{E-43}$$

Now, let us consider the remaining term on the right-hand side of Eq.(E-36). Let $u(x)$ denote the unit step function

$$u(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{E-44}$$

Then the limit of the term in question can be written as

$$\lim_{j\to\infty} \sigma^2 \sum_{\ell=J}^{j-1} \gamma_\ell^2 \, d_\ell^2 \, B_{\ell,j-1} \prod_{m=\ell+1}^{j-1} \lambda_m^2$$

$$= \sigma^2 \lim_{j\to\infty} \sum_{\ell=J}^{\infty} \gamma_\ell^2 \, d_\ell^2 \, B_{\ell,j-1} \, [1 - u(\ell-j)] \prod_{m=\ell+1}^{j-1} \lambda_m^2$$

$$= \sigma^2 \sum_{\ell=J}^{\infty} \gamma_\ell^2 \, d_\ell^2 \lim_{j\to\infty} B_{\ell,j-1} \, [1 - u(\ell-j)] \prod_{m=\ell+1}^{j-1} \lambda_m^2 \tag{E-45}$$

The interchange of limit and sum is justified provided that the sum is absolutely convergent, i.e.,

$$\sum_{\ell=J}^{\infty} \gamma_\ell^2 \, d_\ell^2 < \infty \tag{E-46}$$

Combining Eqs. (E-36), (E-43), and (E-45) at least yields the desired results

$$\lim_{j\to\infty} V_j = 0 \tag{E-47}$$

if Eqs. (E-41), (E-42), and (E-46) are satisfied.

Since, by ratio test, the series $\sum_{j=1}^{\infty} j^{-\alpha}$ diverges for $\alpha \leq 1$ and converges absolutely for $\alpha > 1$, for the ordinary stochastic approximation we require

$$\frac{1}{2} < \alpha \leq 1$$

so that $\sum_{j=1}^{\infty} \gamma_j = \infty$ and $\sum_{j=1}^{\infty} \gamma_j^2 < \infty$ if $\gamma_j = \dfrac{\gamma}{j^\alpha}$.

In the general dynamic case, we require that

$$\sum_{j=1}^{\infty} \gamma_j \, \xi_j = \infty \, , \; \sum_{j=1}^{\infty} \gamma_j^2 \, d_j^2 < \infty \, , \text{ and } \prod_{j=1}^{\infty} \lambda_j^2 < \infty \, .$$

As ξ_j is proportional to λ_j^{-1} by Eqs. (E-30) and (E-38), we can let

$$\xi_j = O(j^{\beta}) \tag{E-48}$$

and

$$d_j = O(j^{-\delta}) \tag{E-49}$$

Eqs. (E-41), (E-42), and (E-46) are satisfied if Eq. (E-20) is satisfied. The constraints given by Eq. (E-20) indicate that the sequence $\gamma_j = \frac{\gamma}{j^{\alpha}}$ cannot be chosen arbitrarily. If the optimum set varies too fast ($\delta < 1/2$), the proposed algorithm will fail to track it. Actually the ordinary stochastic approximation method converges at the rate of $O(j^{-\alpha})$, the tracking operation will definitely fail whenever $\delta > \alpha$ or $\beta > \alpha$.

In [36], it is assumed that

$$\underline{\theta}_{j+1} = \underline{\theta}_j \left(1 + \frac{1}{j}\right) = L(\underline{\theta}_j, j) \tag{E-50}$$

Then

$$\underline{G}_j = (\text{grad } L_j^T = \begin{bmatrix} (1 + \frac{1}{j}) & & \\ & (1 + \frac{1}{j}) & \\ 0 & & (1 + \frac{1}{j}) \end{bmatrix} \tag{E-51}$$

$$\lambda_j = 1 + \frac{1}{j} = O(\frac{1}{j^{o}}) \tag{E-52}$$

$$\beta = \delta = 0 \tag{E-53}$$

so that $\frac{1}{2} < \alpha < 1$ remains unchanged as in the stationary case.

It is also to be noted that although $\gamma_j \to 0$ as $j \to \infty$ or the adjustments become smaller as the adaptation process proceeds, the parameters to be estimated still vary in a way similar to the variation of the optimum set because $\underline{c}_{j+1} = L(\underline{c}_j, j)$ for $j \to \infty$.

APPENDIX F

SUMMARY OF THE KALMAN FILTERING TECHNIQUES

In this appendix the formulas for the discrete time optimum-filter solution are given. Detail derivations can readily be found in the literature.

Define the following nomenclatures:

- $\underline{\theta}_j$: system state vector or system parameters
- $\underline{u}_j$: input or control function
- $\underline{v}_j$: white noise
- $\underline{D}_j$: system output vector
- $\underline{A}$: system dynamics function
- $\underline{G}$: input constraints on system state
- $\underline{H}_j$: constraints on observing the state of the system from the system output

If a linear system is characterized by the difference equations

$$\underline{\theta}_{j+1} = \underline{A}\,\underline{\theta}_j + \underline{G}\,\underline{u}_j \quad \text{(F-1)}$$

$$\underline{D}_j = \underline{H}_j\,\underline{\theta}_j + \underline{v}_j \quad \text{(F-2)}$$

together with the statistics:

$$E\{\underline{u}_j\} = E\{\underline{v}_j\} = 0 \quad \text{for all } j \quad \text{(F-3)}$$

$$E\{\underline{u}_k\,\underline{u}_j^T\} = \underline{Q}_k\,\delta_{kj} \quad \text{(F-4)}$$

$$E\{\underline{v}_k\,\underline{v}_j^T\} = \underline{R}_k\,\delta_{kj} \quad \text{(F-5)}$$

$$E\{\underline{u}_j\,\underline{v}_k^T\} = 0 \quad \text{(F-6)}$$

The optimum filter minimizing the performance criterion

$$J_j = ||\underline{H}_j\,\underline{\theta} - \underline{D}_j||^2_{R_j^{-1}} = (\underline{H}_j\,\underline{\theta} - \underline{D}_j)^T\,\underline{R}_j^{-1}\,(\underline{H}_j\,\underline{\theta} - D_j) \quad \text{(F-7)}$$

is described by

$$\hat{\underline{\theta}}_{j+1}\ \hat{\underline{\theta}}_j + \underline{K}_j\ (d_{j+1} - \underline{h}_{j+1}\ \underline{A}\ \hat{\underline{\theta}}_j) \tag{F-8}$$

where d_{j+1} = the (j+1)-st column of $\underline{D}_{j+1}$

$\underline{h}_{j+1}$ = the (j+1)-st column of $\underline{H}_{j+1}$

and $\underline{K}_j$ is the weighting matrix defined by

$$\underline{K}_j = \underline{P}_{j+1}\ \underline{h}^T_{j+1}\ \underline{R}^{-1}_{j+1} \tag{F-9}$$

$$\underline{P}^{-1}_{j+1} = (\underline{A}\ \underline{P}_j\ \underline{A}^T + \underline{G}\ \underline{Q}_j\ \underline{G}^T)^{-1}$$

$$+ \underline{h}^T_{j+1}\ \underline{R}^{-1}_{j+1}\ \underline{h}_{j+1} \tag{F-10}$$

$\underline{P}_{j+1}$ is the outer product of error of the optimal estimate

$$\underline{P}_{j+1} = E\left\{(\hat{\underline{\theta}}_{j+1} - \underline{\theta}_{j+1})\ (\hat{\underline{\theta}}_{j+1} - \underline{\theta}_{j+1})^T\right\} \tag{F-11}$$

APPENDIX G

DERIVATION OF EQS. (4.3-32) AND (4.3-33)

The general algorithm has been derived by Eq. (4.3-12):

$$\underline{W}_{j+1} = \underline{W}_j + \underline{\Gamma}_{j+1}\, \underline{\eta}_{j+1}(d_{j+1} - \eta^T_{j+1}\, \underline{W}_j) \qquad \text{(G-1)}$$

where

$$\underline{\Gamma}_{j+1} = \frac{1}{\phi}\, \underline{P}_{j+1} \qquad \text{(G-2)}$$

and $\underline{P}^{-1}_{j+1} = (\underline{P}_j + q\, \underline{I})^{-1} + \frac{1}{\phi}\, \underline{\eta}_{j+1}\, \underline{\eta}_{j+1}$ (G-3)

In the stationary case $q = 0$, thus

$$P^{-1}_{j+1} = P^{-1}_j + \frac{1}{\phi}\, \underline{\eta}_{j+1}\, \underline{\eta}^T_{j+1} \qquad \text{(G-4)}$$

or

$$\Gamma^{-1}_{j+1} = \Gamma^{-1}_j + \underline{\eta}_{j+1}\, \underline{\eta}^T_{j+1} \qquad \text{(G-5)}$$

which has a recursive solution

$$\Gamma^{-1}_j = \Gamma^{-1}_o + j\, (\frac{1}{j} \sum_{k=1}^{j} \underline{\eta}_k\, \underline{\eta}^T_k)$$

$$= \Gamma^{-1}_o + j\, \underline{R}_\eta \simeq j\, R_\eta \qquad \text{(G-6)}$$

(derived in Eq. (4.3-23)

Combining Eqs. (G-1) and (G-2) gives

$$\underline{W}_{j+1} = \underline{W}_j + P_{j+1}\, \underline{\eta}_{j+1}\, \phi^{-1}\, (d_{j+1} - \underline{\eta}^T_{j+1}\, \underline{W}_j) \qquad \text{(G-7)}$$

let $P^{-1}_{j+1} = P^{-1}_j + \frac{1}{\phi}\, \underline{\eta}_{j+1}\, \underline{\eta}^T_{j+1}$

be the weighting matrix for the stationary case

and

$$\underline{P}'_{j+1} \triangleq \underline{P}_{j+1} + \underline{B}_{j+1} \qquad \text{(G-8)}$$

be the weighting matrix for the nonstationary case such that

$$(P_{j+1} + B_{j+1})^{-1} = (P_j + q\,\underline{I})^{-1} + \frac{1}{\phi}\,\underline{n}_{j+1}\,\underline{n}_{j+1}^T \tag{G-9}$$

Since from Eq. (G-6)

$$P_j^{-1} \approx 0(j) \tag{G-10}$$

we can write for large j

$$\frac{1}{\phi}\,\underline{n}_{j+1}\,\underline{n}_{j+1}^T << (P_{j+1} + B_{j+1}) \tag{G-11}$$

$$\frac{1}{\phi}\,\underline{n}_{j+1}\,\underline{n}_{j+1}^T << (P_j + q\,\underline{I})^{-1} \tag{G-12}$$

Thus, Eq. (G- 9) becomes

$$(P_{j+1} + B_{j+1})^{-1} \approx (P_j + q\,I)^{-1} \tag{G-13}$$

Since $P_j \approx 0(j^{-1})$ and $B_j \rightarrow B = \text{const.}$, we can write

$$P_j << B_j \quad \text{for large } j$$

or from Eq.

$$B \approx q \tag{G-14}$$

which is Eq. (4.3-32).

Returning to Eq. (G-7) since for large j

$$P'_{j+1} = P_j + B_j \approx B_j \approx B \approx q \,, \tag{G-15}$$

the corresponding algorithm becomes

$$\underline{W}_{j+1} = \underline{W}_j + \frac{q}{\phi}\,\underline{n}_{j+1}\,(d_j - \underline{n}_{j+1}^T\,\underline{W}_j) \tag{G-16}$$

Comparing with the ordinary algorithm

$$\underline{W}_{j+1} = \underline{W}_j + \gamma_j\,\underline{n}_{j+1}\,(d_j - \underline{n}_{j+1}^T\,\underline{W}_j) \tag{G-17}$$

we can replace γ_j by γ_j' such that

$$\gamma_j' = \gamma_j + B = \gamma_j + \frac{q}{\phi} \tag{G-18}$$

which is Eq. (4.3-33) which has $B = 0$ for $q = 0$.

REFERENCES

1. Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series, John Wiley and Sons, N. Y., 1949.

2. Kalman, R. E., and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," J. Basic Engineering, Vol. 83D, 1961.

3. Faran, J. J. and Hills, R., "Wide-Band Directivity of Receiving Arrays," Harvard Univ. Acoust. Res. Lab. Tech. Memo. No. 31, May 1953.

4. Mermoz, H., "Adaptive Filtering and Optimal Utilization of an Antenna," Bureau of Ships Translation No. 903, Oct. 1965.

5. Burg, J., "Three-Dimensional Filtering with an Array of Seismometers," Geophysics, vol. 29, no. 5, pp. 693-713, Oct. 1964.

6. Bryn, F., "Optimum Signal Processing of Three-Dimensional Arrays Operating on Gaussian Signals and Noise," J. Acoust. Soc. Am. vol. 34, no. 3, pp. 289-297, March 1962.

7. Edelblute, D. J., Fisk, J. M., and Kinnison, G. L., "Criteria for Optimum-Signal-Detection Theory for Arrays," J. Acoust. Soc. Am. vol. 41, no. 1, pp. 199-205, Jan. 1967.

8. Schultheiss, P. M., and Tuteur, F. B., "Optimum and Suboptimum Detection of Directional Gaussian Signals in an Isotropic Gaussian Noise Field, Part I - Likelihood Ratio and Power Detector, Part II, Degradation of Detectability Due to Clipping, "IEEE Trans. on Military Electronics, pp. 197-211, July-Oct. 1965.

9. Tuteur, F. B., "On the Detection of Transiting Broadband Targets in Noise of Uncertain Level, "IEEE Trans. on Comm. Tech., vol. Com-15, no. 1, pp. 61-69, Feb. 1967.

10. Tuteur, F. B., "Some Aspects of the Detectability of Broadband Sonar Signals by Nondirectional Passive Hydrophones," The Rand Corporation Research Memo. Rm-4578-ARPA, June 1965.

11. Widrow, B., "Adaptive Sampled Data Systems," International Federation of Automatic Control, Moscow, 1960.

12. Gabor, D., et al., "A Universal Non-linear Predictor, Filter, and Simulator, "Proc. of IEEE (London), Part B, vol. 108, July 1960.

13. Bucy, R. S. and Follin, J. W., "Adaptive Finite Time Filtering," IRE Trans. on Automatic Control, vol. AC-7, July 1962.

14. Narendra, K. S. and McBride, L., "Multiparameter Self-Optimizing Systems Using Correlation Techniques," IEEE Trans. on Automatic Control, Vol. AC-9, Jan. 1964.

15. Widrow, B., "Adaptive Filters I-Fundamentals," Tech. Rept. No. 6764-6, Stanford University, Dec. 1966.

16. Widrow, B., et al., "Adaptive Antenna Systems," Proc. of IEEE, vol. 55, no. 12, 11. 2143-2158, Dec. 1967.

17. Chang, J. H. and Tuteur, F. B., "Methods of Stochastic Approximation Applied to the Analysis of Adaptive Delay Line Filters," Dunham Laboratory Tech. Rept. CT-17, Yale University, Nov. 1967.

18. Glaser, E. M., "Signal Detection; Adaptive Filters," IRE Trans. on Information Theory, vol. IT-7, pp. 87-98, April 1961.

19. Jakowatz, C. V., Shuey, R. L., and White, G. M., "Adaptive Waveform Reception," Proc. 4th London Symp. on Information Theory, Butlerworths, 1961.

20. Daly, R. F., "Adaptive Binary Detectors," Stanford University Elect. Lab. TR 2003-2, Stanford, Calif., June 1961.

21. Scudder, H. J., "Adaptive Communication Receivers," IEEE Trans. on Information Theory, vol. IT-11, pp. 167-174, April 1965.

22. Peterson, W. W., Birdsall, T. G., and Fox, W. C., "The Theory of Signal Detectability," IRE Trans. on Information Theory, vol. IT-4, Sept. 1954.

23. Cox, H., "Array Processing Against Interference," Unpublished Report, Naval Ship Systems Command, Washington, D.C., Oct. 1967.

24. Knapp, C. H., "Optimum Linear Filtering for Multi-element Arrays," General Dynamics/Electric Boat Division Technical Report U 417-66-031, Nov. 1966.

25. Schultheiss, P. M., "Advanced Topics in Linear Systems," Lecture Notes, Yale University, New Haven, Conn.

26. Churchill, R. V., Fourier Series and Boundary Value Problems, McGraw Hill Book Co., New York, 1963.

27. Narendra, K. S., Gallman, P., and Chang, J. H., "Identification of Nonlinear Systems Using Gradient and Iterative Techniques," Dunham Laboratory Tech. Rept. CT-3, Yale University, Aug. 1966.

28. Robbins, H. and Monro, S., "A Stochastic Approximation Method," Annal.of Math. Stat., vol. 21, 1951.

29. Kiefer, J. and Wolfowity, J., "Stochastic Estimation of the Maximum of a Regression Function," Annal.of Math. Stat., vol. 23, 1952.

30. Blum, J. R., "Multidimensional Stochastic Approximation Methods," Annal. of Math. Stat., vol. 25, 1954.

31. Dvoretzky, A., "On Stochastic Approximation," Proc. Third Berkeley Symp. of Math. Stat. and Prob., Univ. of California Press, 1956.

32. Schultheiss, P. M., "Passive Detection of a Sonar Target in a Background of Ambient Noise and Interference from a Second Target," General Dynamics/Electric Boat Research Progress Reports No. 17, Sept. 1964. Also appeared in JASA, vol. 43, no. 3, 1968.

33. Tuteur, F. B., "The Effect of Noise Anisotropy on Detectability in an Optimum Array Processor, "General Dynamics/ Electric Boat Research, Progress Report No. 33, Sept. 1967.

34. Amari, S., "A Theory of Adaptive Pattern Classifier," IEEE Trans. on Electronic Computers, vol. EC-16, no. 3, pp. 299-307, June 1967.

35. Wozencroft, J. M. and Jacobs, I. M., Principles of Communication Engineering, John Wiley, New York, 1965.

36. Dupac, V., "A Dynamic Stochastic Approximation Method," Ann. Math. Stat., vol. 36, pp. 1695-1702, 1965.

37. Abramson, N. and Braverman, D., "Learning to Recognize Patterns in a Random Environment," IRE Trans. Information Theory, vol. IT-8, pp. s-58 to s-63, Sept. 1962.

38. Sakrison, D. J., "Stochastic Approximation: A Recursive Method for Solving Regression Problems," in Advances in Communication Systems, vol. II, edited by A. V. Balakrishman, Academic Press, New York, 1966.

39. Schultheiss, P. M., "Likelihood Ratio Detection of Gaussian Signals with Noise Varying from Element to Element of the Receiving Array," General Dynamics/Electric Boat Research,Progress Report No. 10, January 1964.

40. McGrath, R. J. and Rideout, V. C., "A Simulator Study of a Two-Parameter Adaptive System," IEEE Trans., AC-6, pp. 35-42, Feb. 1961.

41. Kesten, H., "Accelerated Stochastic Approximation," Am. Math. Stat., vol. 29, pp. 41-59, 1958.

42. Comer, T. R., "Some Stochastic Approximation Procedures for Use in Process Control," Amer. Math. Stat., vol. 35, no. 3, 1964.

43. Bryson,A .E . , and Frazier, M ., "Smoothing for Linear and Nonlinear Dynamic Systems," Proc. Systems Optimization Conf.,Ohio, September 1962.

44. Ho, Y. C., "On the Stochastic Approximation Method and Optimal Filtering Theory," J. of Math. Analysis and Applications, vol. 6, pp. 152-154, 1962.

45. Tuteur, F. B., "The Optimum Detector for Nonisotropic Noise," General Dynamics/Electric Boat Research, Progress Report No. 38, Sept. 1968.

46. Lucky, R. W., "Automatic Equalization for Digital Communication," Bell System Technical Journal, vol. 44, pp. 547-588, April 1965.

47. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill Book Co., N. Y., 1965.

OPTIMUM PASSIVE BEARING ESTIMATION IN A SPATIALLY COHERENT NOISE ENVIRONMENT

by

Verne H. MacDonald

Progress Report No. 40

General Dynamics/Electric Boat Research

(8050-31-55001)

September 1969

DEPARTMENT OF ENGINEERING

AND APPLIED SCIENCE

YALE UNIVERSITY

I Introduction

The present report extends the analytical techniques outlined in Progress Report No. 37 ("Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment," by Verne MacDonald and Peter M. Schultheiss) to the physical situation shown in Figure 1. A linear passive hydrophone array is used to estimate the bearing θ of a distant target in the presence of a distant interfering acoustical source at bearing ϕ. The linear array contains M hydrophones arbitrarily placed at points $(R_1 \ldots, R_m)$ relative to an arbitrary origin. Bearing is measured relative to an axis perpendicular to the array axis. The target signal wavefront and the interference wavefront impinging on the array are assumed to be essentially planar; this assumption implies that the target and interference ranges are much greater than the array length.

We obtain a theoretical lower bound on rms bearing estimation error through the Cramér- Rao inequality, and we compare this bound with the rms error of a modified split beam tracker derived in Progress Report No. 29.

As in Progress Report No. 37, we neglect any inhomogeneities of sound velocity or attenuation in the water. Ambient noise is assumed independent from hydrophone to hydrophone, with equal power and identical spectra at all hydrophones. The results obtained here are intended primarily to show to what extent a distant point-source interference degrades target bearing measurement accuracy, but the results also indicate at least qualitatively

the effect of any directional (coherent) component of noise, irrespective of how this component arises. In this report, the interference bearing is assumed to be known; a future report will treat the problem of an interference of unknown bearing.

The procedure for obtaining the Cramér-Rao lower bound is straightforward, but the details are extremely laborious. For that reason, only selected intermediate results are presented in this report. A cumbersome result is obtained for arbitrary signal, interference, and noise spectra and hydrophone spacing. This result is made somewhat more manageable by the assumptions:

1. ambient noise power much greater than signal power.
2. signal, interference, and noise spectra of identical form.
3. uniform hydrophone spacing.

II Sketch of the Procedure for Finding a Theoretical Lower Bound on Mean Square Bearing Estimation Error

Let $p(\underline{x}|\theta,\phi)$ be the joint probability density of the hydrophone data $\underline{x}$ for specified target bearing θ and interference bearing ϕ. The Cramér-Rao inequality places the following lower bound on the variance of any estimate $\hat{\theta}$ of θ based on $\underline{x}$.[1]

$$(1) \quad \overline{(\hat{\theta}-\bar{\hat{\theta}})^2} \geq \frac{(1+db/d\theta)^2}{\overline{\dfrac{-\partial^2 \log p(\underline{x}\,|\theta,\phi)}{\partial\theta^2}}}$$

Here b is bias, and overbars denote averages with respect to the distribution of $\underline{x}$.

The data $\underline{x}$ may take any legitimate form, such as time samples or expansion coefficients of the waveforms produced by the hydrophones. For analytical convenience, we let $\underline{x}$ be the vector $\underline{F}$ of the complex

coefficients of exponential Fourier series expansions of all hydrophone output waveforms over a time interval (-T/2, T/2). If the output of the ith hydrophone is represented by $f_i(t)$, then the Fourier coefficients take the form

$$(2)\quad F_i(\omega_k) = \int_{-T/2}^{T/2} f_i(t)e^{-j\omega_k t}\,dt \qquad (i=1,\ldots,M)$$

$$(3)\quad \omega_k = \omega_o + 2\pi k/T \qquad (k=1,\ldots,n)$$

We will take ω_o, the lower limit on the processed bandwidth, to be 0, but it need not be. Our data are then arrayed in the Mn - dimensional vector

$$(4)\quad \underline{F} = F_1(\omega_1),\ldots, F_M(\omega_1);\ldots, F_1(\omega_n),\ldots, F_M(\omega_n).$$

If s(t) and i(t) represent the signal and interference components respectively of the waveform at the array origin, and if $n_i(t)$ represents the ambient noise component at the ith hydrophone, then the time waveform at the output of the ith hydrophone is

$$(5)\quad f_i(t) = s(t-\Delta_i) + i(t-\delta_i) + n_i(t), \text{ with}$$

$$(6)\quad \Delta_i = (r_i/c)\sin\theta \qquad \delta_i = (r_i/c)\sin\phi,$$

where c is sound velocity. We assume that s(t), i(t), and all $\{n_i(t)\}$ are stationary zero-mean Gaussian processes, with the result that all components of $\underline{F}$ are stationary zero-mean jointly Gaussian random variables. The joint probability density of $\underline{F}$ can then be written as

---

1 - derived in Appendix A.

$$(7)\quad p(\underline{F}|\theta,\phi) = \frac{1}{(2\pi)^{Mn}\det \underline{R}(\theta,\phi)}\, e^{-\frac{1}{2}\underline{F}^T \underline{R}^{-1}(\theta,\phi)\underline{F}^*},$$

where T and * indicate transposition and complex conjugation, respectively. The elements of the correlation matrix $\underline{R}$ are

$$(8)\quad R_{i,k,j,\ell} = \tfrac{1}{2}\overline{F_i^*(\omega_k)\, F_j(\omega_\ell)}.$$

Appendix B represents $\underline{R}$, det $\underline{R}$ and $\underline{R}^{-1}$ in detail. Once the forms of det $\underline{R}$ and $\underline{R}^{-1}$ are established they can be substituted into the probability density (7) and the Cramér - Rao inequality (1) can be applied to this probability density. As a matter of terminology, one may wish to view $p(\underline{F}|\theta,\phi)$ as being a likelihood function $L(\theta)$ or a conditional likelihood function $L(\theta|\phi)$.

III Results for Arbitrary Spectra and Array Geometry

When the Cramér - Rao inequality (1) is applied to the joint probability density (7) without any new assumptions or approximations, the following general result is obtained after many tedious steps:

$$(9)\quad \overline{(\hat{\theta}-\bar{\hat{\theta}})^2} = \mathrm{var}(\hat{\theta}) \geq (1+db/d\theta)^2 \div$$

$$\sum_{k=1}^{n} \frac{2(S^k)^2\omega_k^2\cos^2\theta}{c^2 D^k}\left\{\frac{2(I^k)^2}{D^k}\left[\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}(x_i-x_j)\sin_{ij}{}^k\right]^2 + \right.$$

$$\frac{N^k+MI^k}{N^k}\left[\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}(r_i-r_j)^2\right]\Bigg|, \quad \frac{-I^k}{N^k}\left[\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\left\{-2\left(\sum_{p=1}^{M} r_p\right)(r_i+r_j)\right.\right. +$$

$$\left.\left. 2Mr_i r_j + 2\left(\sum_{p=1}^{M} r_p^2\right)\right\}\cos_{ij}^k + 2M\left(\sum_{p=1}^{M} r_p^2\right) - 2\left(\sum_{p=1}^{M} r_p^2\right)\right]\Bigg\} \quad ,$$

where

$$\text{(10)} \quad D^k \triangleq (N^k)^2 + MN^k(S^k+I^k) + S^kI^k\left[M(M-1)-2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\cos_{ij}^k\right]$$

$$\text{(11)} \quad \sin_{ij}^k \triangleq \sin\left[\frac{r_i-r_j}{c}\,\omega_k(\sin\theta-\sin\phi)\right]$$

$$\text{(12)} \quad \cos_{ij}^k \triangleq \cos\left[\frac{r_i-r_j}{c}\,\omega_k(\sin\theta-\sin\phi)\right] \quad ,$$

and the details of the notation are:

θ,ϕ : target, interference bearings, respectively

$\{r_i\}$: hydrophone locations

c : sound velocity

k : frequency index $(1,\ldots,n)$

i,j,p,q : hydrophone indices $(1,\ldots,M)$

$(S^k,I^k,N^k) = [S(\omega_k),\ I(\omega_k),N(\omega_k)]$: signal, interference, and noise spectra, resp.

We shall assume that the observation time T is sufficiently large so that negligible error is introduced in (9) by converting the sum over the frequency index to an integral with respect to frequency. If one multiplies all terms in the summation over k in (9) by the factor $(T\Delta\omega/2\pi)$, which equals unity, and then lets $\Delta\omega \rightarrow d\omega$, the result may be written

$$
(13)\quad \overline{(\hat{\theta} - \hat{\hat{\theta}})^2} \geq (1 + db/d\theta)^2 \quad +
$$

$$
\int_{\omega_1}^{\omega_n} \frac{TS^2(\omega)\omega^2\cos^2\theta}{\pi C^2 D(\omega)} \left\{ \frac{2I^2(\omega)}{D(\omega)} \left[\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} (r_i - r_j) \sin_{ij}(\omega) \right]^2 + \right.
$$

$$
\frac{N(\omega)+MI(\omega)}{N(\omega)} \left[\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} (r_i - r_j)^2 \right] - \frac{I(\omega)}{N(\omega)} \left[\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \left\{ -2 \left(\sum_{p=1} r_p \right) (r_i + r_j) + \right. \right.
$$

$$
\left. \left. \left. 2Mr_i r_j + 2\left(\sum_{p=1}^{M} r_p^2 \right) \right\} \cos_{ij}(\omega) + 2M\left(\sum_{p=1}^{M} r_p^2 \right) - 2\left(\sum_{p=1}^{M} r_p \right)^2 \right] \right\} d\omega
$$

IV Results For Weak Signal, Identical Spectra, and Uniform Spacing

In order to obtain a less cumbersome result, we now make the following assumptions :

(14) $N(\omega) >> MS(\omega)$ $\qquad \omega_1 \leq \omega \leq \omega_n$

(15) $\left[S(\omega), I(\omega), N(\omega) \right] = \left[SG(\omega), IG(\omega), NG(\omega) \right], \quad \omega_1 \leq \omega \leq \omega_n$

($G(\omega)$ arbitrary)

(16) $r_i = id$ $\qquad i=1,\ldots,M$

Assumption (14) states that the coherent sum of signal power from all the hydrophones is still much smaller than the ambient noise power at any one hydrophone. This assumption permits one to neglect the frequency-dependent part of the factor D^k(10). Condition (15) requires that the signal, interference, and ambient noise processes have identically shaped spectra over the processed frequency band. This assumption is intended primarily to simplify the complicated integral in (13), but it is actually realistic for certain cases of interest. If the target and interference are similar vessels, for instance, then

the spectra of the broad band acoustical waveforms emanating from them may be quite similar. The ambient noise may also be similar in character to the signal and interference processes, except that the noise is likely to have a broader bandwidth. The assumption (15) is realistic if the ratios $S(\omega)/I(\omega)$, $S(\omega)/N(\omega)$, and $N(\omega)/I(\omega)$ remain close to some constant values throughout the processed frequency band. The integral in (13) is simplified considerably, since one in effect replaces the spectra $[S(\omega),I(\omega),N(\omega)]$ by the constants (S,I,N), respectively. Uniform hydrophone spacing (16) offers a typical example of linear array configurations, and it converts the complicated sums involving the $\{r_i\}$ into polynomials in M.

The substitution of assumptions (14)-(16) into the general result (13), together with the assumption that $\omega_1 \neq 0$ and the definition $\omega_{max} \triangleq \omega_n$, yields the following result :

$$(17) \quad \overline{(\hat{\theta}-\hat{\tilde{\theta}})^2} \geq \frac{(1+db/d\theta)^2 \, \pi \, c^2 D}{TS^2 \, \omega_{max}^3 \, d^2 \cos^2\theta} \div$$

$$\left\{ \frac{2I^2}{D} \left\{ \underbrace{\frac{M^5-M}{180}}_{\triangleq\, G_o} - \underbrace{\frac{1}{y^3} \cdot \frac{1}{4} \sum_{p=1}^{M-1} (M-p)^2 \left(y \cos(2py) + \left(py^2 - \frac{1}{2p}\right) \sin(2py)\right)}_{\triangleq\, G_1} \right. \right.$$

$$\underbrace{+ \frac{1}{y^3} \cdot \sum_{p=1}^{M-2} \sum_{q=p+1}^{M-1} \frac{(M-p)(M-q)pq}{(p-q)^2} \left[2y \cos[(p-q)y] + \left((p-q)y^2 - \frac{2}{(p-q)}\right) \sin[(p-q)y]\right]}_{=\, G_2}$$

$$-\frac{1}{y^3}\cdot\underbrace{\sum_{p=1}^{M-2}\sum_{q=p+1}^{M-1}\frac{(M-p)(M-q)pq}{(p+q)^2}\left[2y\cos[(p+q)y]+\left\{(p+q)y^2-\frac{2}{(p+q)}\right\}\sin[(p+q)y]\right]}_{\triangleq G_3}\Bigg\}$$

$$+\Bigg\{\underbrace{\left[1+(M-2)\frac{I}{N}\right]\frac{M^4-M^2}{36}}_{\triangleq G_4}$$

$$-\frac{I}{N}\cdot\frac{1}{y^3}\cdot\frac{M}{3}\underbrace{\sum_{p=1}^{M-1}\left[\frac{M^3-2pM^2-M+p+p^3}{p^2}\right]\left[2y\cos(py)+(py^2-\frac{2}{p})\sin(py)\right]}_{\triangleq G_5}\Bigg\}\Bigg\}$$

$$=\frac{\dfrac{(1+db/d\theta)^2\,\pi c^2 D}{TS^2\omega_{max}^3 d^2\cos^2\theta}}{\dfrac{2I^2}{D}\left[G_0-\dfrac{1}{y^3}(G_1-G_2+G_3)\right]+\left[1+(M-2)\dfrac{I}{N}\right]G_4-\dfrac{1}{N}\left[\dfrac{1}{y^3}G_5\right]}$$

where

(18) $D \triangleq N^2 + MN(S+I) + M(M-1)SI$

(19) $y \triangleq \frac{d\,\omega_{max}}{c}(\sin\theta - \sin\phi)$.

The result (17) can be made more comprehensible by considering its asymptotic forms for large y (interference remote in bearing from target) and small y (interference near target in bearing.

A. Remote Interference

If $y >> 1$,* the oscillatory expressions G_1, G_2, and G_3 may be neglected with respect to G_o; and G_5 may be neglected with respect to $(M-2)G_4$, for $M >> 1$. The lower bound given by (17) is then approximately

$$(20)\quad \overline{(\hat{\theta}-\bar{\hat{\theta}})^2} \geq \frac{\dfrac{(1+db/d\theta)^2\,\pi\,c^2 D}{TS^2\omega_{max}^3\,d^2\cos^2\theta}}{\dfrac{I^2}{D}\cdot\dfrac{M^5-M}{90}+\left[1+(M-2)\dfrac{I}{N}\right]\dfrac{M^4-M^2}{36}}$$

Further approximations can be made in this result if one assumes either that ambient noise dominates interference or vice-versa. The factor D(18) takes different forms in these two limiting cases.

$$(21)\quad D \cong \begin{cases} N^2 + MN(S+I) & MI << N \\ MIN + M(M-1)IS & MI >> N \end{cases}$$

Substituting (21) into (20) yields

$$(22)\quad \overline{(\hat{\theta}-\bar{\hat{\theta}})^2} \geq \begin{cases} \dfrac{(1+db/d\theta)^2\,36\,\pi\,c^2}{T\omega_{max}^3\,d^2\cos^2\theta(M^4-M^2)\left[1+(M-2)\dfrac{I}{N}\right]}\cdot\dfrac{N^2+MSN}{S^2} & (MI<<N) \\ \dfrac{(1+db/d\theta)^2\,36\pi\,c^2}{T\omega_{max}^3\,d^2\cos^2\theta\left[(M^4-M^2)-\dfrac{8}{5}M^3+2M\right]}\cdot\dfrac{N^2+(M-1)SN}{S^2} & (MI>>N) \end{cases}$$

---

* example : d=2ft., c = 5,000 ft./sec., ω_{max} = $2\pi \times 5{,}000$, $(\sin\theta-\sin\phi)=1$;

then $y = (d/c)\,\omega_{max}\,(\sin\theta-\sin\phi) = 4\pi$

If one sets I/N equal to zero in the noise-dominant version of the above inequality, one obtains a lower bound on the variance of bearing estimates in the absence of interference. Note that the lower bound in the interference-dominated case is almost identical to the lower bound in the no-interference case. This condition obtains, of course, only when the interference is remote in bearing from the target. When the number of hydrophones is large, the effect of a remote interference is equivalent to the loss of 2/5 of a hydrophone *. Note that we have assumed that the strength of the interference has negligible effect on the bias $b(\theta)$.

B. Near Interference

If $y \lesssim 1/M$,** one may replace the oscillatory functions in (17) by the first two terms of their respective Maclaurin series. When the target and interference are very close in bearing,

$$(23) \qquad \overline{(\hat{\theta}-\hat{\theta})^2} \gtrsim \frac{\dfrac{(1 + db/d\theta)^2 \, 36 \, \pi \, c^2 D}{TS^2 \, \omega_{max}^3 \, d^2 \cos^2 \theta}}{(M^4-M^2)\left[1 + (M^4-M^2)\frac{I}{N} \cdot \frac{y^2}{10}\right]}.$$

---

*see Appendix C for explanation.

**example : d= 1ft., c = 5,000 ft./sec., $\omega_{max} = 2\pi \times 5{,}000$, $(\sin\theta - \sin\phi) = .003$, M= 50; then $y = (d/c)\,\omega_{max}(\sin\theta - \sin\phi) = 6\pi \times 10^{-3} < 1/50 = .02$

To find the results for the noise-dominant and interference-dominant cases, we substitute the expressions for D in (21) into (23):

$$(24)\ \overline{(\hat{\theta}\ \hat{\theta})^2} \geq \begin{cases} \dfrac{(1 + db/d\theta)^2\ 36\ \pi\ c^2}{T\omega_{max}^3\ d^2\cos^2\theta(M^4-M^2)\left[1 + (M^4-M^2)\dfrac{I^2}{N^2}\cdot\dfrac{y^2}{10}\right]}\cdot\dfrac{N^2+MSN}{S^2} & (MI<<N) \\ \dfrac{(1 + db/d\theta)^2\ 360\ \pi\ c^2}{T\omega_{max}^3\ d^2\cos^2\theta(M^4-M^2)(M^2-1)}\cdot\dfrac{[N+(M-1)S]^2}{S^2}\cdot\dfrac{1}{y^2} & (I>>N\ ,\ y>>\dfrac{3}{M^2}\cdot\dfrac{N}{I}) \\ \dfrac{(1 + db/d\theta)^2\ 36\pi\ c^2\ M}{T\omega_{max}^3 d^2\cos^2\theta(M^4-M^2)}\cdot\dfrac{I[N+(M-1)S]}{S^2} & (MI>>N\ ,\ y<<\dfrac{3}{M^2}\ \dfrac{N}{I}) \end{cases}$$

By setting (I/N) equal to zero in the first line of (24), we obtain a lower bound on the variance of $\hat{\theta}$ in the absence of interference. The last line sets a lower bound on the variance of $\hat{\theta}$ when the target and interference are essentially coincident in bearing. A strong interference at the same bearing as the target is seen to increase the variance of $\hat{\theta}$ by a factor of approximately (MI/N).

The weak interference versions of (22) and (24) closely resemble the no-interference result given by Eq.(35) of Progress Report No.37. The primary object of including the weak-interference results in (22) and (24) is to indicate the first-order effect of an interference on target bearing estimation accuracy. The situation which presents considerable practical difficulties is that in which a strong interference is present.

C. Strong Interference

By substituting the strong-interference version of D(21) into the general result (17), one obtains the lower bound for the strong-interference case with arbitrary y :

$$(25)\quad \overline{(\hat{\theta}-\bar{\hat{\theta}})^2} \ \geq \ \frac{\dfrac{(1+db/d\theta)^2 \pi c^2}{T\omega_{max}^3 d^2 \cos^2\theta} \cdot M\dfrac{N^2}{S^2}}{\dfrac{2}{M} G_o - \dfrac{1}{y^3}(G_1 - G_2 + G_3) + (M-2)G_4 - \dfrac{1}{y^3}G_5} .$$

V Split-Beam Tracker Performance

Progress Report No.29 presents results for the variance of bearing estimates obtained in the presence of a single point-source interference with a modified version of the split-beam tracker. We shall repeat some of those results, which are valid for the following conditions :

1. signal weak with respect to ambient noise.
2. uniform hydrophone spacing.
3. signal, interference, and ambient noise spectra flat over the processed band.

In the split-beam tracker, the estimate $\hat{\theta}^*$ is obtained by varying the steering angle θ_o to determine for what value of θ_o the output z equals zero. This value of θ_o is then taken as $\hat{\theta}$. The variance of z, σ_z^2, can be derived as a function of θ, θ_o, ϕ, the array geometry, and the spectra of the signal, interference, and noise processes. Then the variance of $\hat{\theta}$, $\sigma_{\hat{\theta}}^2$, is given approximately by the formula

$$(26)\quad \sigma_{\hat{\theta}}^2 \simeq \left. \frac{\sigma_z^2}{|\partial\bar{z}/\partial\theta_o|} \right|_{\theta_o = \theta}$$

This formula is valid if $\partial\bar{z}/\partial\theta_o$ is essentially constant for θ_o in the interval $(\theta-\sigma_{\hat{\theta}}, \theta+\sigma_{\hat{\theta}})$.

From equations (38-42,86,90) of Progress Report No.29 we have the the following results (for white or prewhitened spectra):

* The notation used here differs somewhat from that used in Progress Report No.29.

(27) $$\sigma_{\bar{z}}^2 \Big|_{\theta_o = \theta} =$$

$$\frac{4\pi\, N^2}{T\omega_{max}^{\,3}} \left[(M'-1)^2 + \frac{1}{2}(M'-2)^2 - 2(M'-1)(M'-2)\frac{\sin y}{y} + \frac{1}{2}(M'-2)^2 \frac{\sin 2y}{2y} \right]$$

(28) $$\partial \bar{z} / \partial \theta_o \Big|_{\theta_o = \theta} = \omega_{max}\, S\, \frac{d}{c} \cos\theta \left[1 - 2\,\frac{\sin y}{y} + 2\,\frac{1-\cos y}{y^2} \right] M'(M'-1)^2,$$

where M' is simply $M/2$. Substituting (27) and (28) into (26), we obtain

(29) $$\sigma_{\hat{\theta}}^2 \cong \frac{64\pi c^2}{T\omega_{max}^{\,3}\; d^2 \cos^2\theta M^2 (M-2)^4} \cdot \frac{N^2}{S^2} \times$$

$$\frac{\left[(M-2)^2 + \tfrac{1}{2}(M-4)^2 - 2(M-2)(M-4)\frac{\sin y}{y} + \frac{(M-4)^2}{4} \cdot \frac{\sin 2y}{y} \right]}{\left[1-2\,\frac{\sin y}{y} + 2\,\frac{1-\cos y}{y^2} \right]^2}$$

The simpler asymptotic expressions for large and small y are

(30) $$\sigma_{\hat{\theta}}^2 \cong \begin{cases} \dfrac{96\,\pi c^2}{T\omega_{max}^{\,3}\; d^2 \cos^2\theta M^2 (M-2)^2} \cdot \dfrac{N^2}{S^2} & (y \gg 1,\ M \gg 1) \\[2ex] \dfrac{(64)^2 \pi\, c^2}{T\omega_{max}^{\,3}\; d^2 \cos^2\theta M^2 (M-2)^4} \cdot \dfrac{1}{y^4} \cdot \dfrac{N^2}{S^2} & (y \ll 1),\ M \gg 1) \end{cases}$$

The latter expression becomes unbounded as y approaches zero, that is, as the target and interference bearings converge.

Below some critical value of y, depending on all the parameters in (30), including θ, the expression for $\sigma^2_{\hat{\theta}}$ in (30) becomes invalid because the condition for the validity of (26) is no longer met. The form of the latter expression in (30), however, suggests the very reasonable conclusion that it essentially impossible to estimate a target bearing extremely close to the interference bearing, using the implementation under consideration. It should be noted that (30) becomes invalid for any value of y when $|\cos\theta|$ falls below some critical value, again because the condition on the validity of (26) is not met. As long as neither y nor $|\cos\theta|$ is too small, however, (30) gives an accurate and meaningful result.

VI A Comparison of the Split-Beam Tracker Variance with the Cramér-Rao Lower Bound

Figures 2 through 4 present a graphical comparison of the split-beam tracker variance, according to (29), and the Cramér-Rao lower bound for the strong-interference case according to (25), for various values of M. In plotting the Cramér-Rao lower bound, we assume that $db/d\theta$ is much smaller than unity, so that the factor $(1+ db/d\theta)^2$ in (25) can be ignored. This assumption is invalid for θ near $\pm \pi/2$,* but it should be valid for most of the possible range of θ simply because in a good estimator one would expect bias to be much smaller than a radian for all values of θ. If $b(\theta)$ is a well-behaved function, the condition that $b(\theta)$ is small implies that $db/d\theta$ is also small over most of the allowable range of θ. The lower bound which we have plotted is

(31) $$\sigma_{CR}^2 = \frac{1}{\dfrac{-\partial^2 \log\, p(\underline{F}|\theta,\phi)}{\partial\theta^2}}$$

It should be remembered that the results for the split-beam tracker, designated σ_{sbt}^2, are also invalid for θ near $\pm\,\pi/2$ and for y very near zero, as explained at the end of the previous section.

Figure 4 displays the ratio $\sigma_{sbt}^2 \;/\; \sigma_{CR}^2$ as a function of y for various values of M. In general, this ratio is the complicated quotient of equations (29) and (25), but the form of the ratio is relatively simple for very large or very small y, also assuming a very small signal-to-noise power ratio. From equations (24) and (30) we have

(32) $$\frac{\sigma_{sbt}^2}{\sigma_{CR}^2} \simeq \begin{cases} 2.67\ (1 + 2.4/M) & y \gg 1,\ M \gg 1 \\ \dfrac{11.4}{y^2}\left(1 + \dfrac{8}{M}\right) & y \ll 1,\ M \gg 1 \end{cases}$$

This expression, of course, becomes invalid for very small y, because σ_{sbt}^2 becomes invalid. It does indicate, however, that when the target and interference bearings are very close, the variance of the modified split-beam tracker estimate greatly exceeds the Cramér-Rao lower bound.

The curves for both the split-beam tracker variance (Fig.2) and the Cramér-Rao lower bound (Fig.3) clearly exhibit an overall variation as M^{-4}. For y near zero, the curves rise sharply for both the split-beam tracker and the lower bound.

---

* It happens that $\partial p(\underline{F}|\theta,\phi) \;/\; \partial\theta = 0$ for $\theta \pm \pi/2$. The final paragraph of Appendix A explains that this condition implies that $db/d\theta = -1$ for $\theta = \pm\,\pi/2$.

It is clear in both cases, however, that an increase in the number of hydrophones pushes the "break" value of y progressively closer to zero, in other words, an increasing number of hydrophones should permit a system accurately to measure target bearings progressively closer to the interference bearing.

The shape of the split-beam tracker curves varies only weakly with the number of hydrophones. The lower bound curves, on the other hand, become progressively flatter, beyond the steep rise for small y, as M increases; that is, the performance of an ideal bearing estimator should be essentially independent of the angular separation between target and interference for a large number of hydrophones so long as the separation exceeds some small minimum value. For large values of M, the lower-bound curves can be approximated very well simply by connecting the asymptotic curves for large y and small y.

Fig.4 shows that the ratio of the split-beam tracker variance to the lower bound depends only weakly on the number of hydrophones. It is true, however, that as M increases, the split-beam tracker performance edges slightly closer to the lower bound. Beyond the region of small y, the ratio remains less than about 4 for an array containing 10 or more hydrophones. The performance of the modified split-beam tracker is reasonably good, then, unless target and interference bearings are too close. The minimum angular separtion between target and interference for satisfactory performance can be expressed roughly in terms of the beamwidth of the array. The beamwidth is determined by considering the average signal-derived output $\tilde{z}$ of a conventional detector * as a function of the parameter y', which is defined exactly as y above, except that ϕ is now interprepted to be the steering angle.

The beamwidth is defined in terms of y' as the distance between the two values of y' for which $\bar{z}$ falls to half its maximum value, which occurs at $y' = 0$. As an arbitrary standard of adequate performance one might require that the variance of a practical estimator be no greater than ten times the Cramèr-Rao lower bound. From Fig. 4 one can determine the values of y at which the curves reach the value 10. These values turn out to be approximately equal to the beamwidth of the array, assuming a flat signal spectrum with cutoff frequency ω_{max} . Thus, by this arbitrary standard, the modified spit-beam tracker offers adequate performance as long as signal and interference bearings are separated by at least a beamwidth.

The point at which the Cramèr-Rao curves in Fig. 3 change from steep to relatively flat can also be described quite accurately in terms of beamwidth. The change occurs at approximately 1.75 times the bandwidth.

VII Conclusion

The Cramèr-Rao inequality indicates that the presence of a point-source interference raises the lower bound on target bearing estimation variance over that obtained in the absence of any interference. If the target and interference bearings are separated by an angular difference of more than approximately twice the array beamwidth, the increase in the lower bound is small. In fact, for a large number of hydrophones, the increase is equivalent to the loss of only 2/5 of a hydrophone. If the angular separation between target and interference is on the order of a beamwidth or less, the increase is substantial. When the target and interference bearings coincide (the worst case), the lower bound is

*A conventional detector consists of an array of hydrophones followed by a bank of variable delays for the purpose of steering, then a summer, squarer, and low-pass filter.

MI/N times greater than in the no-interference case. Since in the no-interference case the lower bound varies as M^{-4}, the lower bound for the case of coincident target and interference bearing varies as M^{-3}. In theory, then, even this lower bound can be made arbitrarily small by making M sufficiently large , while all other parameters remain constant.

The modified split-beam tracker discussed in Progress Report No. 29 yields a bearing estimation variance no larger than 4 times the Cramér-Rao lower bound, as long as the angular separation between the target and interference is greater than roughly twice the beamwidth. If the angular separation is substantially smaller than a beamwidth, however, the bearing estimation variance is unsatisfactorily large. If the target and interference bearings coincide, this implementation is completely incapable of measuring target bearing. Thus, although the modified split-beam tracker offers reasonably good performance when the interference is remote in bearing from the target, it is unsatisfactory when the target and interference are very close in bearing, and a different implementation must be sought for this case.

Appendix A: Derivation of the Cramér - Rao Inequality

Begin by writing the following equation, which is in effect a definition of bias $b(\theta)$, as the discrepancy between the mean value of the estimate $\hat{\theta}(\underline{x})$ and the true value θ:

$$\text{(A-1)} \quad \overline{\hat{\theta}(\underline{x})} = \int_{R_{\underline{x}}} \hat{\theta}(x)\, p(\underline{x}|\theta,\phi)\, d\underline{x} = \theta + b(\theta),$$

where $R_{\underline{x}}$ is the domain of $\underline{x}$. Now differentiate both sides of (A-1) with respect to θ:

$$\text{(A-2)} \quad \int_{R_{\underline{x}}} \hat{\theta}(\underline{x}) \frac{\partial p(\underline{x}|\theta,\phi)}{\partial\theta}\, d\underline{x} = 1 + db/d\theta$$

Let $f(\theta)$ be any function of θ which is not a function $\underline{x}$.

$$\text{(A-3)} \int_{R_{\underline{x}}} f(\theta) \frac{\partial p(\underline{x}|\theta,\phi)}{\partial\theta}\, d\underline{x} = f(\theta) \frac{\partial}{\partial\theta} \int_{R_{\underline{x}}} p(\underline{x}|\theta\ \phi) dx = f(\theta)\frac{\partial}{\partial\theta}(1) = 0$$

Subtract (A-3) from (A-1); multiply and divide the integrand by $p(\underline{x}|\theta,\phi)$:

$$\text{(A-4)} \int_{R_{\underline{x}}} \left[\hat{\theta}(\underline{x}) - f(\theta)\right] \frac{\partial p(\underline{x}|\theta,\phi)}{\partial\theta}\, d\underline{x}$$

$$= \int_{R_{\underline{x}}} \left[\hat{\theta}(\underline{x}) - f(\theta)\right] \frac{1}{p(\underline{x}|\theta,\phi)} \frac{\partial p(\underline{x}|\theta,\phi)}{\partial\theta}\, p(\underline{x}|\theta,\phi)\, d\underline{x}$$

$$= \int_{R_{\underline{x}}} \left[\hat{\theta}(\underline{x}) - f(\theta)\right] \frac{\partial \log p(\underline{x}|\theta,\phi)}{\partial\theta}\, p(\underline{x}|\theta\ \phi) d\underline{x} = 1 + db/d\theta$$

The Schwarz inequality reads

$$\text{(A-5)} \int_{R_{\underline{x}}} f^2(\underline{x}) d\underline{x} \int_{R_{\underline{x}}} g^2(\underline{x}) d\underline{x} \geq \left[\int_{R_{\underline{x}}} f(\underline{x}) g(\underline{x}) d\underline{x}\right]^2$$

Let $f(\underline{x}) \triangleq \left[\hat{\theta}(\underline{x}) - f(\theta)\right]\sqrt{p(\underline{x}|\theta,\phi)}$ and

$g(\underline{x}) \triangleq \dfrac{\partial \log p(\underline{x}|\phi,\theta)}{\partial\theta}\sqrt{p(\underline{x}|\theta,\phi)}$ Apply (A-5) to (A-4):

$$\text{(A-6)} \quad \int_{R_{\underline{x}}} \left[\hat{\theta}(x) - f(\theta)\right]^2 p(\underline{x}|\theta,\phi)d\underline{x} \int_{R_{\underline{x}}} \left[\frac{\partial \log p(\underline{x}|\theta,\phi)}{\partial\theta}\right]^2 p(\underline{x}|\theta\ \phi)d\underline{x}$$

$$\geq (1 + db/d\theta)^2$$

This inequality is equivalent to

$$\text{(A-7)} \quad \overline{[\hat{\theta}(x)-f(\theta)]^2} \geq \frac{(1 + db/d\theta)^2}{\overline{\left[\dfrac{\partial \log p(\underline{x}|\theta\ \phi)}{\partial\theta}\right]^2}}$$

Equation (A-7) holds for any $f(\theta)$. Using variational techniques, one can show that the choice for $f(\theta)$ which minimizes the left side of (A-7) is $f(\theta) = \bar{\hat{\theta}}$. Thus for arbitrary $f(\theta)$,

$$\text{(A-8)} \quad \overline{[\hat{\theta}(x)-f(\theta)]^2} \geq \overline{[\hat{\theta}(x)-\bar{\hat{\theta}}]^2} \geq \frac{(1 + db/d\theta)^2}{\overline{\left[\dfrac{\partial \log p(\underline{x}|\theta,\phi)}{\partial\theta}\right]^2}}$$

The right side of (A-7) is often derived as a lower bound on mean square error $(f(\theta)=\theta)$, and so it is. It should be emphasized, however, that if bias is present, mean square error cannot achieve this lower bound. A tighter lower bound on mean square error is as follows:

$$\text{(A-9)} \quad \overline{(\hat{\theta}-\theta)^2} = b^2(\theta)+\overline{(\hat{\theta}-\bar{\hat{\theta}})^2} \geq b^2(\theta) + \frac{(1+db/d\theta)^2}{\overline{\left[\dfrac{\partial \log p(\underline{x}|\theta\ \phi)}{\partial\theta}\right]^2}}$$

The rightmost two members of (A-9) are seen to be equivalent to (1) through the following identity, which is proved in various textbooks:[1]

1. For example, Harry L. Van Trees, Detection, Estimation, and Modulation Theory, Part 1, Section 2.4.

$$\text{(A-10)} \quad \overline{\left[\frac{\partial \log p(x|\theta,\phi)}{\partial\theta}\right]^2} = -\overline{\frac{\partial^2 \log p(x|\theta,\phi)}{\partial\theta^2}}\ .$$

A point deserving comment is the question of what properties, if any are common to the bias functions associated with all possible estimators $\hat{\theta}(\underline{x})$ which might be used in the same situation. The derivation of the Cramer-Rao inequality indicates one fact about the derivative $db/d\theta$. Suppose that the derivative $\partial p(\underline{x}|\theta,\phi)/\partial\theta$ equals zero identically for some value of θ, which we shall designate as θ_o, independently of $\underline{x}$. According to (A-2), then, unless $\hat{\theta}(\underline{x})$ is infinite, $db/d\theta$ must equal -1 for θ equal to θ_o, so that both sides of (A-2) equal zero. Since we know that θ is restricted to a finite interval, the possibility of an infinite $\hat{\theta}(x)$ is unacceptable. The only conclusion that can be drawn, therefore, is that $db/d\theta$ must be -1 at the point θ_o for any acceptable estimator $\theta(x)$. This conclusion may be repeated symbolically as

$$\text{(A-11)} \quad \left[\left.\frac{\partial p(x|\theta,\phi)}{\partial\theta}\right|_{\theta=\theta_o} = 0\right] \rightarrow \left[\left.\frac{db}{d\theta}\right|_{\theta=\theta_o} = -1\right]$$

Note that the behavior of $db/d\theta$ away from the point θ_o and the behavior of $b(\theta)$ at all points are in no way specified by (A-11). These depend on the specific form of $\hat{\theta}(\underline{x})$.

Appendix B: The Correlation Matrix $\underline{R}$, its Determinant, and its Inverse

We shall state the form of the correlation matrix, its determinant, and its inverse without proof. The vector $\underline{F}$ is written at the edges of $\underline{R}$ to indicate which two elements of $\underline{F}$ correspond to each element of $\underline{R}$.

$$
\text{(B-1)}\quad \underline{R}=T/2
\begin{array}{c}
\left[\begin{array}{cccc|c|c}
S^1+I^1+N^1 & a^1_{12}S^1+b^1_{12}I^1 & & \cdots a^1_{1M}S^1+b^1_{1M}I^1 & & \\
a^1_{21}S^1+b^1_{21}I^1 & S^1+I^1+N^1 & & \cdots a^1_{2M}S^1+b^1_{2M}I^1 & \cdots & \text{all elements} \\
\vdots & \vdots & & \vdots & & \\
a^1_{M1}S^1+b^1_{M1}I^1 & a^1_{M2}S^1+b^1_{M2}I^1 & & \cdots S^1+I^1+N^1 & & \text{zero} \\
\hline
 & \vdots & & & \ddots & \vdots \\
\hline
 & & & & & S^n+I^n+N^n \cdots \\
 & \text{all elements zero} & & & \cdots & \vdots \quad \ddots
\end{array}\right]
\\
\begin{array}{cccccc}
F_1(w_1) & F_2(w_1) & \cdots F_M(w_1) & \cdots & F_1(w_n)\cdots & F_M(w_n)
\end{array}
\end{array}
\begin{array}{c}
F_1(w_1) \\ F_2(w_1) \\ \vdots \\ F_M(w_1) \\ \vdots \\ F_1(w_a) \\ \vdots
\end{array}
$$

$a^k_{ij} \triangleq e^{jw_k(\Delta i-\Delta j)}$ $\qquad$ $b^k_{ij} \triangleq e^{jw_k(\delta_i-\delta_j)}$

$\Delta_i \triangleq (r_i/c)\sin\theta$ $\qquad$ $\delta_i \triangleq (r_i/c)\sin\phi$

$(S^k, I^k, N^k) \triangleq [S(w_k),\ I(w_k),\ N(w_k),]$

θ,ϕ: target, interference bearings, respectively

$S(w)$, $I(w)$, $N(w)$: signal, interference, noise spectra, respectively

T: observation time

$\{r_i\}$: hydrophone locations

c: sound velocity

i,j: hydrophone indices (1,..,M)

k: frequency index (1,...,n)

Both $\underline{R}$ and $\underline{R}^{-1}$ have the property that only the elements of the MxM diagonal submatrices are nonzero. Each of the n diagonal submatrices corresponds to a different frequency but has the same form.

$$\text{(B-2)} \quad \det \underline{R} = (T/2)^{Mn} \prod_{k=1}^{n} (N^k)^{M-2} \{(N^k)^2 + MN^k(S^k+I^k) + S^k I^k [M(M-1) - 2\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \cos_{ij}^{k}]\}$$

where

$$\text{(B-3)} \quad \cos_{ij}^{k} \triangleq \cos\left[\frac{r_i - r_j}{c}\,\omega_k(\sin\theta - \sin\phi)\right]$$

In the absence of interference, det $\underline{R}$ is independent of θ, but with interference present, it depends on both θ and ϕ.

The elements of $\underline{R}^{-1}$ are as follows, with i,j hydrophone indices, and k,ℓ frequency indices:

$$\text{(B-4)} \quad R^{-1}_{i,k;j,\ell} = \frac{\partial\delta_{k}\cdot(N^k)^{M-3}}{T \det \underline{R}} \times$$

$$\begin{cases} (N^k)^2 + (M-1)N^k(S^k+I^k) + S^K I^k [(M-1)(M-2) - 2\sum\limits_{\substack{p=1 \\ p,q\neq i}}^{M-} \sum\limits_{q=p+1}^{M} \cos_{pq}^{k}] & (i=j); \\ N^k(a_{ij}^k S^k + b_{ij}^k I^k) = S^k I^k [M-2)(a_{ij}^k + b_{ij}^k) - \sum\limits_{\substack{p=1 \\ p\neq i,j}}^{M} (a_{pj}^k b_{ip}^k + a_{ip}^k b_{pj}^k)] & (i\neq j), \end{cases}$$

where

$$\text{(B-5)} \quad \delta_{k\ell} = \begin{cases} 1 & k=\ell \\ 0 & k\neq\ell \end{cases}$$

Apprndix C: Equivalent Cost in Hydrophones of a Remote Interference

The two limiting forms of the result (22) indicate that the factor (M^4-M^2) in the no-interference case (I/N=0) changes to $[M^4-M^2-(8/5)M^3+2M]$ in the remote-interference case. The equivalent cost x in hydrophones is found by substituting (M-x) for M in the no-interference expression and setting it equal to the remote-interference expression.

$$\text{(C-1)} \quad [(M-x)^4-(M-x)^2] = M^4-M^2-(8/5)M^3+2M$$

$$M^4-4M^3x \quad +\ldots \; -(M^2 +\ldots) = M^4-M^2 - (8/5)M^3 +\ldots$$

$x = 2/5 \quad$ for $M >> 1$

Thus, having a strong remote interference with M hydrophones is in a sense equivalent to having no interference and (M - 2/5) hydrophones.

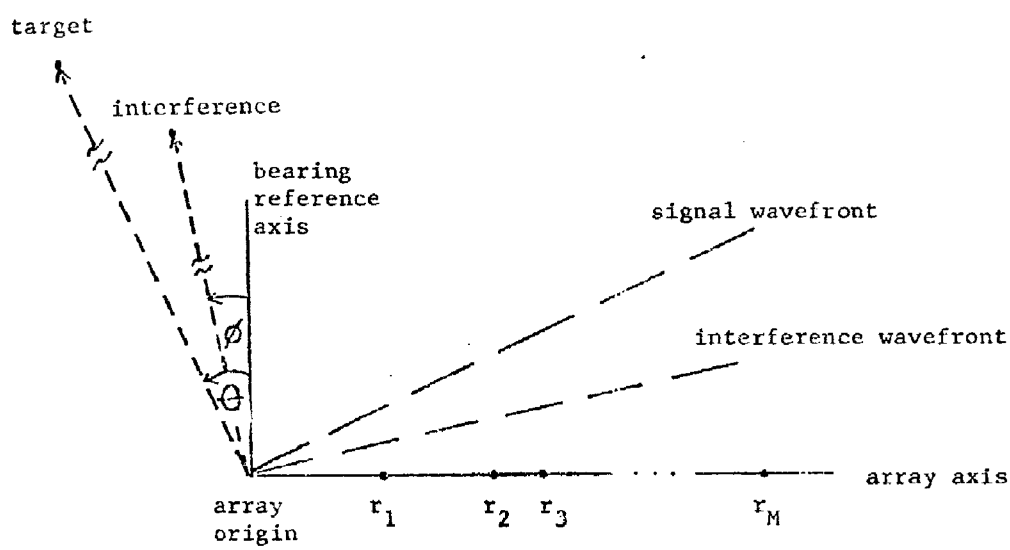
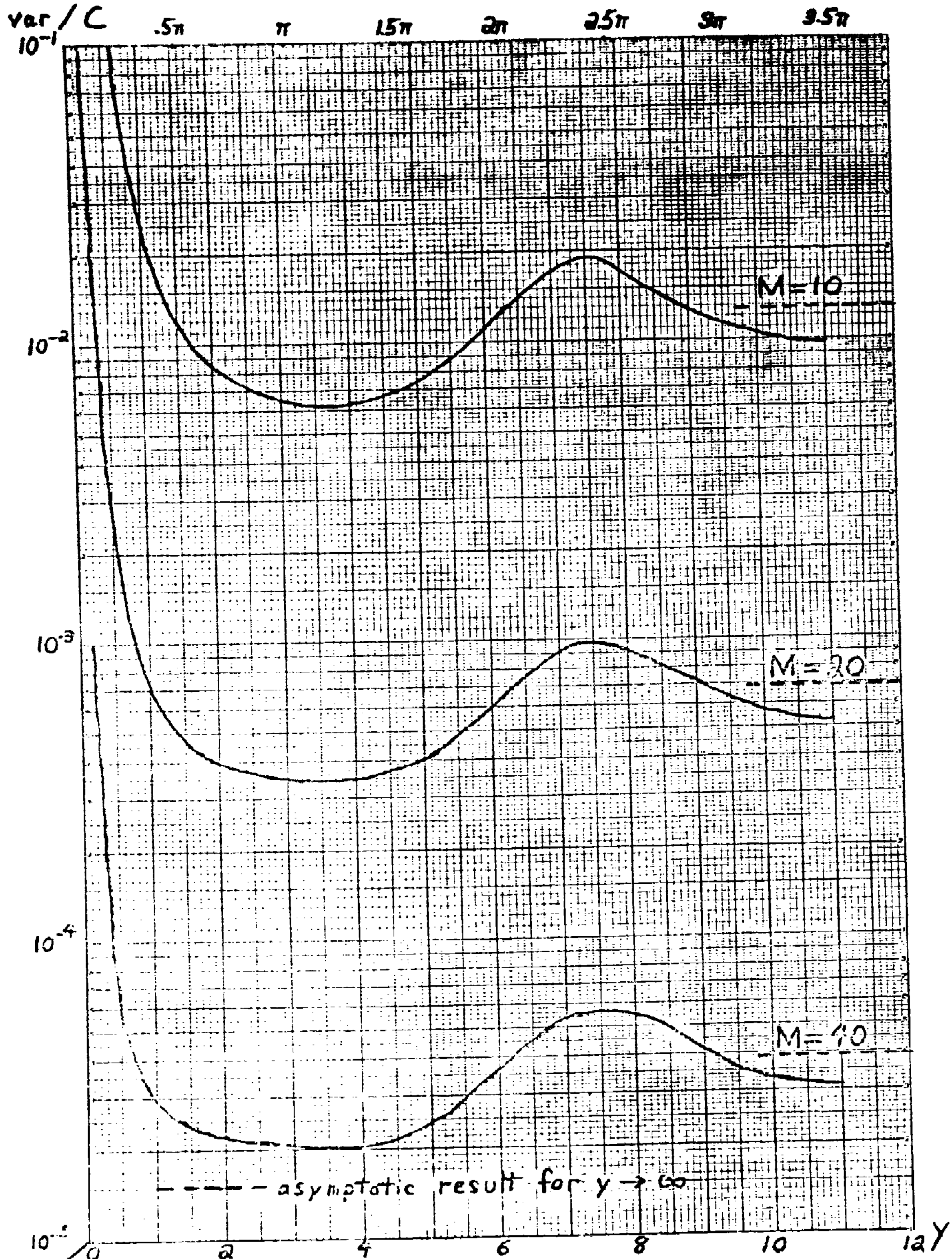
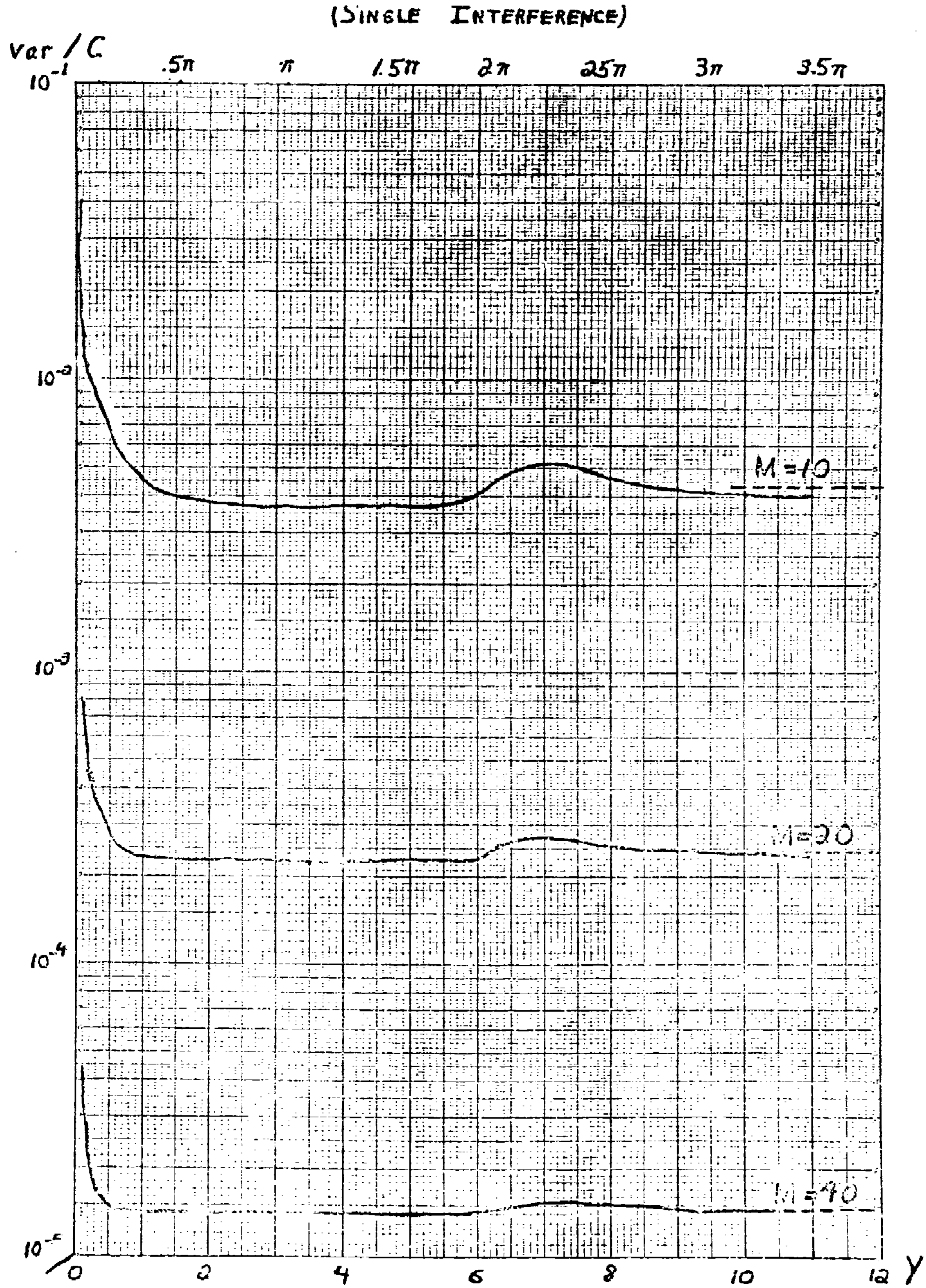


Figure 1. Array Geometry

(SINGLE INTERFERENCE)



$C \triangleq \dfrac{\pi c^2 N^2/S^2}{T\omega_{max}^3 d^2 \cos^2\theta}$ $\qquad y \triangleq \omega_{max}\dfrac{d}{c}(\sin\theta - \sin\phi)$


(SINGLE INTERFERENCE)
Var / C
.5π
π
1.5π
2π
2.5π
3π
3.5π
10^-1
10^-2
10^-3
10^-4
10^-5
M=10
M=20
M=40
0
2
4
6
8
10
12
Y

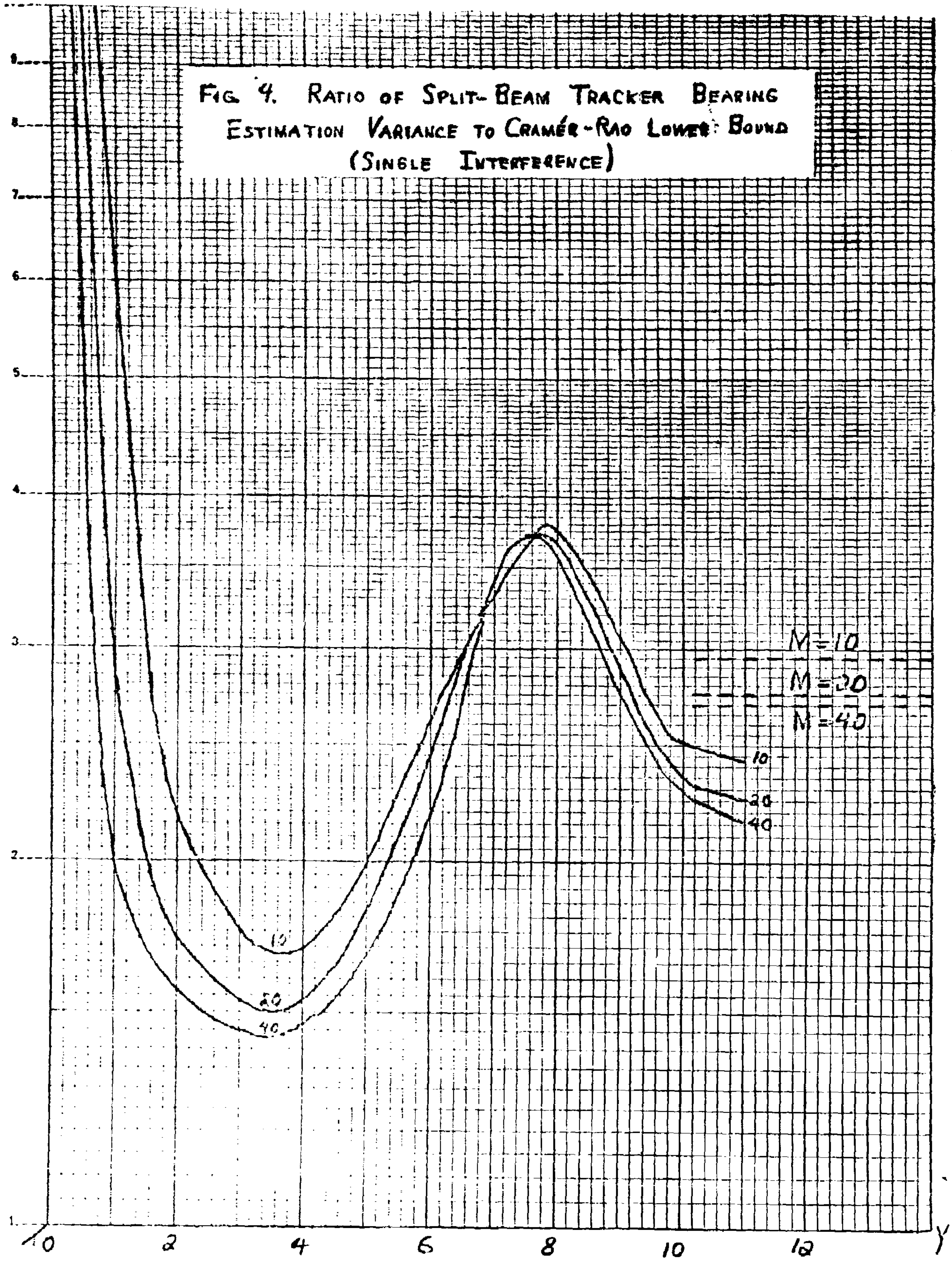


Fig. 4. Ratio of Split-Beam Tracker Bearing Estimation Variance to Cramér-Rao Lower Bound (Single Interference)
M=10
M=20
M=40
10
20
40
0
2
4
6
8
10
12
Y
1
2
3
4
5
6
7
8
9

SPACE-TIME PROPERTIES OF SONAR DETECTION MODELS

by

James Peyton Gray

Progress Report No. 41

General Dynamics/Electric Boat Research

April 1970

DEPARTMENT OF ENGINEERING
AND APPLIED SCIENCE

YALE UNIVERSITY

SUMMARY

Space-time Properties of Sonar Detection Models

James Peyton Gray

April 1970

A measure theoretic structure that is general enough to encompass most models of signal detection is used to investigate singularities in models of sonar detection. Singularities that appeared in previous sonar work are shown to derive from simplified modeling of sound generation and transmission. The existence of singularities in a model of sonar detection is also shown to seriously restrict the usefulness of that model in investigations of sonar array design. Finally, methods for avoiding singularities are discussed.

CONTENTS

1. INTRODUCTION

The problem addressed in this dissertation did not just materialize out of thin air, but rather evolved during an investigation of algorithms for sonar array design. In this chapter the sonar array design problem will be analyzed and the relevance of later chapters established. The link between the analysis and the subsequent work is the question of singularity of models of passive sonar detection as the number of hydrophones increases without bound. In the limit, this means continuous observation.

1.1 Sonar Array Design

Passive sonar detection systems extend from sensors (hydrophones used as acoustic energy to electric energy transducers) to outputs which range from simple audio for human interpretation to complex situation displays of diverse kinds. Too complex to be designed as a whole, these systems must be partitioned; the traditional partitioning allows a single lead out of the array subsystem into the post-array processor and assumes that the spatial processing is to be done in the array subsystem and time processing of the resulting signal is to be done by the post-array processor. This yields a factored sonar detection system in the sense of Middleton [1] (Figure 1.1). That is, there are two operators: one, the array processor, is a function of hydrophone position and signal location, but is independent of signal and noise statistics (the signal is assumed to be a stationary point source). The time operator is

dependent upon signal and noise statistics, but is independent of the array geometry and target location. This factorization leads to advantages in design, construction and operation that are manifest. In many cases these advantages more than offset the sub-optimal performance of the factored system.

Under these conditions, designing the array means placing the hydrophones and combining their outputs to maximize some figure of merit at the array subsystem output. Array output SNR is often used, with delayed summing, i.e., beam-forming, as the processing. Alternatively, if the signal is confined to a narrow frequency band, beam width or the ratio of main beam level to maximum sidelobe level may be used. Although the cost is increased considerably, the beam width and side lobe levels can be partially controlled by introducing shading factors. Other variations in the array processing that have been investigated include multiplying the outputs of several phones together before delayed summing (Shearman [1]) and hard limiting the signals before delayed summing (Anderson [1], Usher [1], Schultheiss [1]).

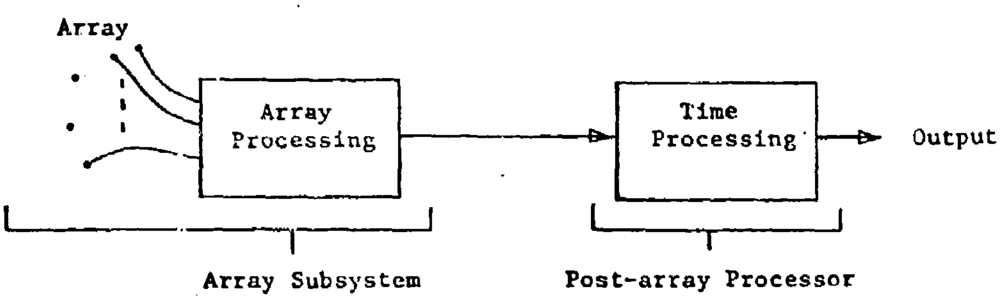


Figure 1.1 A Factored Detection System

Within this basic beam-forming approach, relatively little attention has been paid to the question of hydrophone location. For instance, Skolnik, et al [1] used dynamic programming (Bellman [1]) (even though the principle of optimality does not hold) to position hydrophones in a linear array for "best" beam patterns. Their paper contains references to related design approaches. The more usual design method is to assume an ad hoc array geometry and then to compute beam patterns and shading factors (Lowenstein [1]), sometimes with explicit inclusion of inter-phone acoustic coupling and the directionality of each phone. This computation is so involved that an optimization algorithm (for phone position) based upon it is impractical.

Even if the computation involved were reasonable, conventional beam pattern optimization suffers from a serious defect: the beam-forming/time processing type of detector is not optimal. In fact, the optimal detector factors into separate space and time processors only under very restrictive assumptions: if the spatial processing is to be done first as in conventional beam-forming, a strong signal assumption must be valid (Middleton and Groginsky [1]). If the detector is to be optimum for small signals then a portion of the time processing must be performed first (Goode [1], Bryn [1]), and the observation time must be long. This fact is especially pertinent today, when the optimal processor can be implemented digitally as a special purpose computer, for, there is no reason to believe that an array designed for beam-forming is the best array for use with an optimal processor. Consider, for instance, a noise source located within the volume available to the array: a beam-forming

algorithm would put all the phones at a distance from the noise source, while an optimal processor would place one phone right on top of the local noise and then subtract the local noise from the other phones.

One man has considered array design with optimal processing; N.T. Gaarder, in his dissertation and two closely related papers (Gaarder [1][2][3]) has found optimal radii for circular point detector arrays when a liklihood ratio processor is used on the array outputs. Unfortunately, his results depend in an essential way upon a trick evaluation of the eigenvectors of the covariance of the noise.# For our purposes, however, the chief defect in Gaarder's work is the assumption of isotropic noise; an assumption which he considers essential to the analysis (Gaarder [2] page 48).

This trick is the same one used earlier by Vanderkulk [1] and depends on the point detectors being arranged at the k roots of -1 in the complex plane. Gaarder does not seem to have been aware of either Vanderkulk [1] or Bryn [1], at least he does not reference them in Gaarder [2] or [3].

1.2 Array Optimization Algorithms

Why has Gaarder been the only one to consider array design with optimal processing? Simply because any general approach must founder on the shoals of numerical analysis. In order to illuminate this point, we will try to formulate an array optimization algorithm for liklihood ratio processing of the array outputs.

Let the observed pressure field be $\xi(x_j,t)=\xi_j(t)$ at each of k point sampling hydrophones where x_j is a vector in 3-space. A set of n linear functionals $\{f_i\}$ can be applied to $\xi_j(t)$ to derive a set of observation coefficients. These can be arranged in a single vector η:

$$\eta_m=(\xi_j,f_i) \qquad m=i+n(j-1)$$

We can form the likli-hood ratio:

$$\Lambda(\eta)=\mathrm{Prob}(\eta\varepsilon\{\mathrm{signal+noise}\})/\mathrm{Prob}(\eta\varepsilon\{\mathrm{noise\ alone}\})$$

Assuming that signal present and signal absent are a priori equally probable, assuming small SNR, and assuming independent Gaussian signal and noise, a well known computation yields

$$\Lambda'=2\ln\Lambda(\eta)=(\eta,R^{-1}QR^{-1}\eta)$$

where Q and R are the covariances of the signal and noise processes, respectively. Λ' is a function of the n·k observations η_m, but it is also a function of the 3k coordinates

$$x=\{x_j\} \qquad j=1...k$$

For small signals at the input, the output signal to noise ratio is a measure of the detector's performance. Some algebra gives

$$SNR_k(x)=\mathrm{Tr}(R^{-1}QR^{-1}Q)/((\mathrm{Tr}R^{-1}Q)^2+2\mathrm{Tr}(R^{-1}QR^{-1}Q))$$

This expression reveals the essential difficulty that any optimization algorithm must overcome: the independent variables, $x=\{x_j\}$ $j=1,\ldots,k$ enter the function $SNR_k(x)$ through a matrix inversion. This means that the principle of optimality does not hold, so that a simultaneous optimization in $3 \cdot k$ variables is necessary instead of the k optimizations in 3 variables that could be handled easily. Needless to say, an analytic derivation of the extreme points is very difficult, and has been carried out only by Gaarder [2] and only in a very special case.

The computation of $SNR_k(x)$ at a single point x is a formidable task, especially since the values of k which are of practical interest are in the 100+ range. Since the inversion of a matrix much larger than 20 by 20 is generally conceded to be possible only after extensive study of the particular case,* a straight forward computational approach would be a massive undertaking. Prodigious amounts of machine time and several man-years of effort could be consumed with no guarantee of success except for small k.

* Wilkinson [1]. Matrix inversion algorithms fail on matrices that are ill-conditioned; correlation matrices are generally ill-conditioned. Extensive study of a given matrix would be needed to estimate the degree of ill-conditioning and adapt an inversion algorithm to it. Special techniques, such as the use of programmed multiple-precision arithmetic might be needed to do the job. All of this implicitly assumes long detection times so that the linear functionals can be Fourier transform coefficients. In this special case, R will consist of n uncoupled k by k submatrices on the main diagonal leading to a much easier inversion of R than is the case for a general R

Hamming once said "The purpose of computing is insight..." (Hamming [1]). At the very least, then, before making an investment of the magnitude estimated above, one would like to have some hope of a substantial reward of insight. In this case, one would like to know that the optimized array would offer substantially improved performance. This means knowing that

$$\Delta_k = \max_x \mathrm{SNR}_k(x) - \mathrm{SNR}_k(x_o)$$

is at least 5 or 6 db, where x_o represents any reasonable array geometry-uniformly distributed hydrophones for instance.

It is clear that $\max_x \mathrm{SNR}_k(x)$ is a strictly increasing* function of k. One way, then, to estimate Δ_k for a particular signal and noise distribution, would be to substitute $\lim_k \max_x \mathrm{SNR}_k$ in place of $\max_x \mathrm{SNR}_k$. Unfortunately, the limiting value is less available than the maximum. If the limit exists, though, Δ_k can be estimated thusly: taking L>>k,

$$\Delta_k = \mathrm{SNR}_L(x_o) - \mathrm{SNR}_k(x_o) + \Delta_L + \varepsilon_{kL}$$

$$\Delta_L = \max_x \mathrm{SNR}_L(x) - \mathrm{SNR}_L(x_o)$$

$$\varepsilon_{kL} = \max_x \mathrm{SNR}_k(x) - \max_x \mathrm{SNR}_L(x)$$

where ε_{kL} increases to 0 as k and L go to infinity and Δ_L decreases to 0 as L goes to infinity. For fixed k and large enough L, then, Δ_k is approximated by

$$\Delta_k' = \Delta_k - \Delta_L - \varepsilon_{kL} = \mathrm{SNR}_L(x_o) - \mathrm{SNR}_k(x_o)$$

* Let $\underline{x}$ be the point at which $\mathrm{SNR}_{k-1}(x)$ attains its maximum. Add an additional hydrophone at $x_k \neq x_j$ for all $x_j \in \underline{x}$. Since it picks up some signal power, its output improves the performance of the optimal detector, so that there exists x^* for which

$$\max_x \mathrm{SNR}_k(x) >= \mathrm{SNR}_k(x^*) > \max_x \mathrm{SNR}_{k-1}(x)$$

with an error which is positive for large enough L. This means that if we compute Δ_k' and it is small, then there surely is not enough reward to justify a large investment in array design, at least for that particular signal and noise model. On the other hand, if Δ_k' is large, then we can not be certain of a substantial reward, but we can be hopeful. In words, this estimate is derived by comparing the detection performance of systems with moderately dense and highly dense hydrophone arrays, where the arrays are both in the same volume.

The underpinning of this heuristic calculation is the existence of $\lim_k \max_x SNR_k(x)$ *. Or, put another way, the model must be non-singular in the limit of continuous observation. This first step is not trivial; many models of passive sonar detection are singular in the limit of continuous observation. The body of this dissertation is devoted to understanding why this is so, and determining how to avoid it.

* The particular expression given for SNR(x) is valid only for small signal at the output of the detector. The difficulty in evaluating R^{-1} still remains in any figure of merit for the optimal detector, however.

1.3 Previous Work in Singularity of Models

In the last section we saw that if a model is singular (i.e., if detection is perfect) in the limit of continuous observation then it will not be useful in array design. We are led by this route to consider singularities in detection models.

A model is a mere semblence, a mathematization of a portion of reality. As a construct it can best be judged by its fruitfulness; as an image of the real it must be judged by the faithfulness of its representation. Now, singularity as a property of the mathematical model is neither good nor bad but merely interesting. When considered as a reflection of reality, however, it is an effront to our sensibilities - experience teaches that nowhere is there perfection, everything is fuzzy around the edges, nothing works perfectly. Therefore, a singular model can not be a faithful representation of the real.

Of course, singularity is a fascinating subject in its own right, but, to deepen our understanding of the world we need models of greater faithfulness, which in detection theory means non-singular ones. The tension implicit in this statement is reflected in the literature dealing with singularity: it has been studied by two nearly disjoint sets of workers. On the one hand stand the mathematicians and closely related types (e.g. Yaglom [1]). Generally speaking this group has concerned itself with conditions for orthogonality or equivalence of Gaussian measures (so-called structural questions). This problem was solved to a purist's satisfaction by Feldman [1], who proved that two Gaussian measures are either orthogonal or equivalent. This result was proved, at

least partially, at about the same time by Hajek [1]. A less general form of this theorem had been proved earlier by Grenander [1]. Restated and clarified by Kallianpur and Oodaira [1] with the aid of the theory of reproducing kernel Hilbert spaces, (Aronszajn [1][2]) the essential ideas as they relate to singularity [illegible] models of detection are discussed in section 3.2.

Recognizing that Feldman's conditions are too abstract for easy application, a slightly more practical group of mathematicians has tried to derive different formulations of Feldman's theorem. Conditions in terms of the covariances of the Gaussian processes defined by the measures would be infinitely more useful. Feldman [3], Hinsel: and Varberg [1] have managed to do this for special cases. Feldman's results are extensions of ones provided by Slepian [1] when one of the processes has a rational spectral density. Shepp [1] shows that two zero mean Gaussian processes [illegible] interval that [illegible] type covariances are equivalent if and only if the [illegible] slopes of the covariances at zero are equal. Many authors have written on some aspect of this question, e.g. Shepp [2] [3], Rao [1].

All of the [illegible] far as [illegible] defined on [illegible] functions of a single real variable. Analog[illegible] and [illegible] this [illegible] mat[illegible]

[illegible] singular [illegible]

With the exception of Slepian [1] (and he is not an engineer by training) the singularities have variously: been accepted without question (Martel and Mathews [1]); been considered relatively unimportant (Vanderkulk [1], Gaarder [2][3]); been eliminated by addition of a white self-noise at the detector (Root [1]). The white noise solution to the singularity dilema works well enough when functions of a single variable are under consideration, but, as both Vanderkulk and Gaarder show, array based detection models can be singular as k increases to infinity even in the presence of white noise at each point detector, although detector performance, as measured by the array gain, does increase extremely slowly with k. This is a puzzling behavior, and although it raises serious questions about the adequacy of the models that are being manipulated*, neither author offers a discussion or rationalization.

There has not been, then, a satisfactory treatment of model singularity, especially for non-gaussian problems and for sonar models. The remaining chapters offer a treatment of singularity in models of detection and communication which is useful in sonar and which is independent of the random processes involved.

* If detection becomes perfect as the number of phones increases, how can one be sure that the model is close to the real world, unless a comparison with a better model has established a range of validity for the simpler one?

2. MODELS OF COMMUNICATION AND DETECTION SYSTEMS

That system analyses are performed upon models of reality is obvious. It is less so that problems more difficult than performance evaluation deal not with models but with classes of models. For example, questions of system sensitivity force consideration of all those models which are close, in some metric, to a given one, while system synthesis means attempted maximization of a performance criterion over a collection of models. These facts make an explicit discussion of classes of models desireable.

The basic theme of this chapter is the introduction of classes of models, which is accomplished in sections 2.1 and 2.2. Some basic mathematical problems are discussed in section 2.3 and then two examples are presented in section 2.4. In 2.5, several methods of topologizing a class of models are discussed. In section 2.7 common performance criteria are derived and applied in an example. Finally, singular models are defined and their effects evaluated in section 2.8. The final section, 2.9, is an example of singularity sneaking into a non-factorable model

A moderate background in the measure theoretic development of probability theory is required for this chapter. Knowledge of Halmos [1] or Kingman [1] is sufficient. In addition, the development of stochastic processes as probability measures on appropriate spaces of sample functions is assumed. The first and last chapters of Parthasarathy [1] contain relevant material. Notation is standard, or is defined as it appears. Distribution is used synonymously

with probability measure. Operator, map and transformation are used to mean function. Supp(μ) is any support of the measure μ, while $\overline{\text{supp}}(\mu)$ is the closed support of μ, that is, the smallest of all of the closed sets which support μ.

2.1 The Concept of Classes of Models

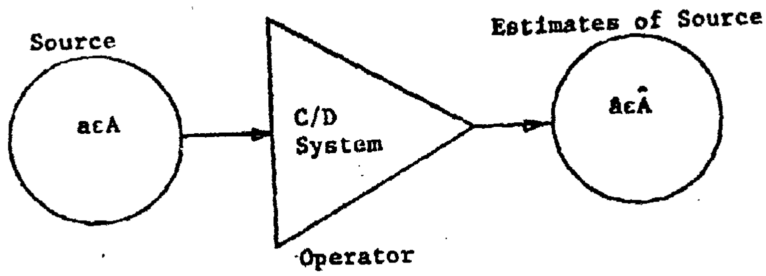
By suppressing all detail, a communication or detection (C/D) system can be modeled by an operator which maps a set of source messages into a set of estimated source messages. See Figure 2.1. But, it is the detail which is of interest to the analyst; this internal structure may be modeled conveniently by a series of composed operators, as in the example displayed in Figure 2.2. Models of particular C/D systems may require different operators, but the general structure shown is sufficient to represent all open-loop C/D systems.

Since C/D systems are probabilistic in nature, the operators must be stochastically determined; this is depicted in Figure 2.3. The model, taken as a whole, must then be a probability space, but, one with a rather complicated internal structure. A point in this space for the example consists of a single source character, a, and four operators, g, n, t, e, so that $\hat{a}$, the estimated signal, is related to a by

$$\hat{a}=(g \circ n \circ t \circ e)(a)$$

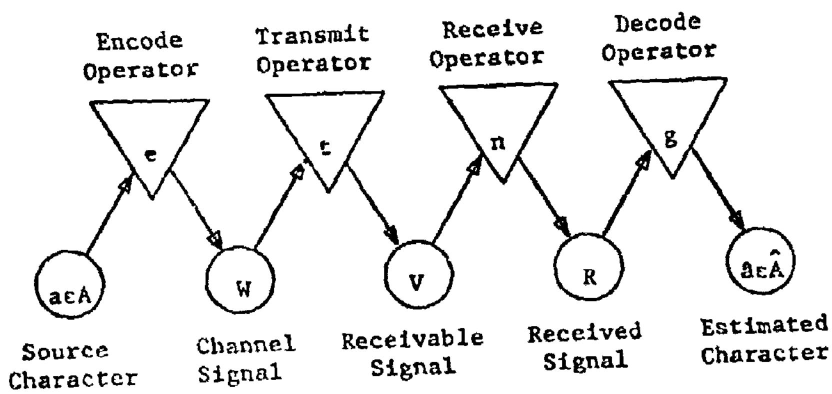
Another point might consist of the same original message a, followed by a different set of operators, g_1, n_1, t_1, e_1, so that

$$\hat{a}_1=(g_1 \circ n_1 \circ t_1 \circ e_1)(a)$$



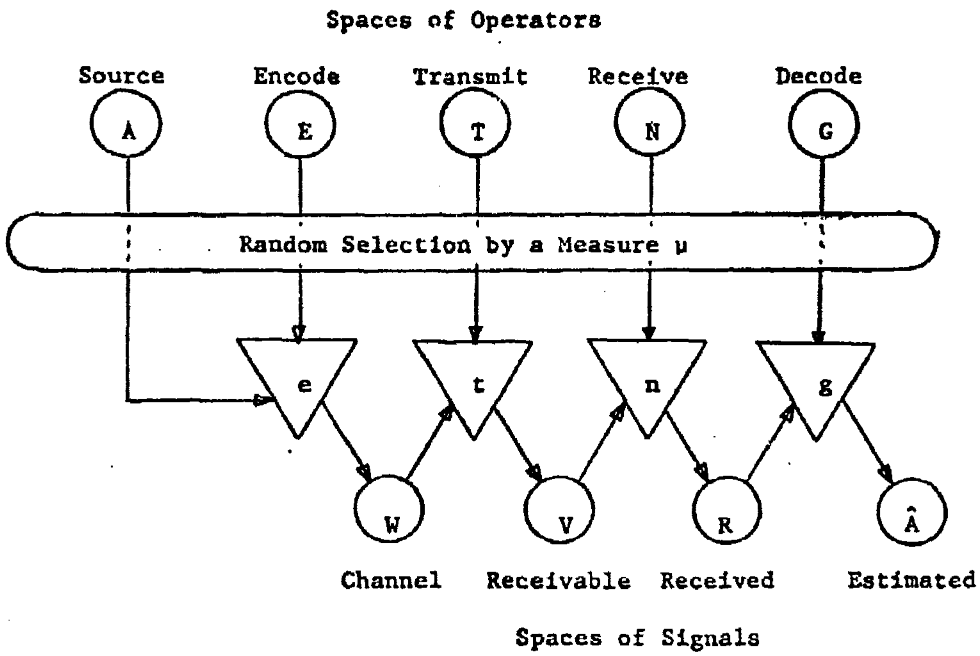
Communication/Detection Systems

Figure 2.1



More Detailed Description of a C/D System
Functioning upon a Message Character a

Figure 2.2



General Model of C/D Systems with Explicit
Representation of Randomness

Figure 2.3

This representation of C/D models leads naturally to consideration of classes of models. By limiting membership in the sets A, W, V, R, E, T, N and G, a particular class of models can be constructed which embodies the constraints imposed on the C/D system by the outside world (in jargon, by the supra-system of which the C/D system is a component). Different measures μ then represent different models. Questions of optimization, sensitivity and the like, which imply a reference class (find the optimal system in this class of systems) can now be discussed with the universe of permissable systems explicitly represented.

2.2 Definitions

Having motivated everyone, it is time to be more precise. We begin by formalizing the discussion given in the preceeding section. Let S_i be a complete metric space with metric $d_i(\cdot,\cdot)$ and obtain a measurable space, also denoted by S_i where no confusion can result, by generating a σ-field B_i from the open sets in S_i. Elements of B_i are Borel sets of S_i and B_i is the Borel σ-field of S_i.

If S_{2k} is a space of measurable mappings of S_{2k-1} into S_{2k+1} for k=1,2,3, ... L and μ is a probability measure on the measurable product space

$$S=S_1 \times S_2 \times S_4 \times \ldots \times S_{2L} = \prod_{j\varepsilon E} S_j$$

where

$$E=\{1\} \text{ union } \{2,4,6,\ldots 2L\}$$

then the 2L+2 tuple

$$M=(S_1,S_2,\ldots,S_{2L+1},\mu)$$

is called a level L C/D model. It will also be called an L-stage (C/D) model.

That this definition is not vacuous is demonstrated by the following example. Let $S_1=S_3=(\underline{R},B)$, that is, the real line under the standard metric. Elements of S_2 will be additive translations: if $t_a \varepsilon S_2$ then

$$t_a:S_1 \rightarrow S_1+a=S_3 \qquad \text{for all } a\varepsilon\underline{R}$$

S_2 becomes isometric (and isomorphic with respect to addition, although that does not concern us here) to S_1 and S_3 if we take

$$d_2(t_a,t_b)=|a-b|$$

hence S_1, S_2 and S_3 are all complete separable metric spaces. Since elements of S_2 are clearly measurable mappings, every 4-tuple $M=(S_1,S_2,S_3,\mu)$ where μ is any measure on $S=S_1xS_2$ is a level 1 C/D model or a one stage C/D model.

For odd $j>1$, the measure μ should in some sense induce a measure, call it μ_j, on S_j. Let $C(T_i)$ be a measurable cylinder* set in S with base T_i in S_i. Define (for even i only)

$$\mu_i(T_i)\equiv\mu(C(T_i))$$

so that μ_i is the marginal measure on the measurable space S_i. Now extend this notation to prism sets*. If F is any subset of indices from the set E, then $C(T_F)$ is the prism set defined by

$$C(T_F)\equiv \prod_{j\varepsilon F} C(T_j)$$

* Cylinder set as defined in Cramer [1], page 17. Prism set is his rectangle set when the index set F contains only two indices.

where the $C(T_j)$ are all cylinder sets. Obviously,

$$\mu_F(T_F) \equiv \mu(C(T_F))$$

defines a marginal measure on the measurable product space $\prod_{j \in F} S_j$

Taking

$$E_{2n} = \{1,2,4,\ldots,2n\} \qquad \text{for all } n>1$$

we are now able to define an induced measure for odd k by:

$$\mu_k(Q) \equiv \int_{\substack{\prod S_j \\ j \in E_{k-1}}} \chi_{t_1}(t_2^{-1} \circ t_4^{-1} \circ \ldots \circ t_{k-1}^{-1} Q) \mu_{E_{k-1}}(dt_1 dt_2 \ldots dt_{k-1}) \qquad \text{(E2.1)}$$

where $\chi_{t_1}(Z)$ is the indicator function of the set Z (that is, $\chi_t(Z)=1$ if $t \in Z$, and 0 otherwise). This definition holds whenever the integral exists. If the integral fails to exist, μ_k is undefined. C/D models of such a pathological nature are interesting in the same way that the plague is: both are to be avoided. There are interesting mathematical problems here; they are discussed in section 2.3.

An important simplification in this relationship occurs when μ is a product of marginal measures, that is, when the stochastic operations of the model are independent. In that case,

$$\mu = \mu_1 \mu_2 \mu_4 \ldots \mu_{2L} = \prod_{j \in E} \mu_j \qquad \text{(E2.2)}$$

and

$$\mu_k(Q) = \int_{S_{k-1}} \mu_{k-2}(t^{-1} Q) \mu_{k-1}(dt) \qquad \text{for all k odd>1}$$

A C/D model with this property will be termed factorable.

When the target space S_k is $\underset{\sim}{R}^n$, the distribution is equivalent to a density m_k (allowing the density to be a generalized function):

$$m_k(x) = \int_{S_{k-1}} m_{k-2}(t^{-1} x) \mu_{k-1}(dt) \qquad \text{(E2.3)}$$

If μ_{k-1} has an associated density:

$$\mu_k(Q)= \int_{S_{k-1}} \mu_{k-2}(t^{-1}Q)m_{k-1}(t)dt \tag{E2.4}$$

and if both m_{k-2} and m_{k-1} exist:

$$m_k(x)= \int_{S_{k-1}} m_{k-2}(t^{-1}x)m_{k-1}(t)dt \tag{E2.5}$$

Conditional probability calculations play an important role in detection theory, so it will be well to consider conditioning of C/D models. Suppose that a particular character has been transmitted, and define

$$E^*_{2n}=\{2,4,\ldots 2n\}$$

$$S^*_{2n}= \prod_{j\varepsilon E^*_{2n}} S_j$$

and let $\underline{t}$ be elements of S^*_{2n}. If $\mu^*= \int_{S_1}\mu$ we see that

$$\mu^a_{k+1}(Q)= \int_{S^*_k} \chi_a(\underline{t}^{-1}Q)\mu^*_k(d\underline{t})$$

is the measure induced on S_{k+1} conditioned by transmission of a. Extended to set conditioning: (for $\mu_1(R)>0$)

$$\mu^R_{k+1}(Q)=(\mu_1(R))^{-1} \int_R \mu^a_{k+1}(Q)\mu_1(da)$$

2.3 Induced Measures*

In section 2.2, C/D models were defined which consist of an original space of randomly chosen characters and a number of spaces of stochastic operators which take the source characters through a series of transformations. Now we want to deal with the measures induced on the spaces S_k, for odd $k>1$. Under easily satisfied conditions, the measures defined by E2.1 will be induced, that is, the integral will exist and so will the induced measure defined by that integral.

The most expeditious approach to this subject requires a certain redirection of our thought. To begin, we note that an L-stage model

$$M_L=(S_1,S_2,\ldots,S_{2L+1},\mu)$$

is more than an arbitrary 2L+2 tuple. An assumed structure exists: every even index space S_k consists of operators which map S_{k-1} into S_{k+1}, or put another way, for $k=2j$, $j=1,2,\ldots,L$ there is a map Φ_k of (S_{k-1},S_k) into S_{k+1} which is defined by

$$\Phi_k:(u,v)\to v(u)\epsilon S_{k+1} \qquad \text{for all } (u,v)\epsilon(S_{k-1},S_k)$$

The set $\{\Phi_{2j}:j=1,2,\ldots,L\}$, one map Φ_k for each stage of the model, represents the totality of ways in which source characters are transformed into detected characters. By modifying the definition of Φ_k slightly, this can be made more explicit. Let Θ_k map (S,S_{k-1}) into (S,S_{k+1}) according to the rule:**

* This section has benefited greatly from conversations with Prof. M. Keane of the Yale Mathematics Department.

** We remember that $S=(S_1,S_2,S_4,\ldots,S_{2L})$

$$\Theta_k(S,S_{k-1})=\Theta'_k(S,S_{k-1},S_k)$$
$$=(S,\Phi_k(S_{k-1},S_k))$$
$$=(S,S_{k+1})$$

In every essential way, then, Θ_k and Φ_k are equivalent. In particular, Θ_k is a measurable map iff Φ_k is.

Going one step further, let's compose several Θ_k starting at the first stage, k=2:

$$\Psi_k=\Theta_k o\Theta_{k-2}o\ldots o\Theta_4 o\Theta_2$$

We see that Ψ_k maps (S,S_1) into (S,S_{k+1}) for each even k up to 2L. Ψ_k is measurable if each Φ_j, j<=k is measurable, since composition preserves measurability. Ψ_k is not in exactly the form we would like, however. To get that form, let ν be the binary selection operator: if x is any n-tuple, $x=(x_1,x_2,\ldots,x_n)$ where n may be infinity, then

$$j\nu x\equiv x_j$$

Now define Γ_k as

$$\Gamma_k(S)=2\nu\Psi_k(S,S_1)=2\nu(S,S_{k+1})=S_{k+1}$$

Γ_k expresses the way in which S_{k+1} is mapped into by the model. Note that Γ_k is measurable iff Ψ_k is, with the result that Γ_k is measurable whenever Φ_j is for all j<=k. For reference, we state this as a theorem:

THEOREM 1 Γ_k is measurable if Φ_j is measurable for all j<=k

The existence of induced measures on S_{k+1} can now be expressed as:

THEOREM 2 If Γ_k is measurable then the integral in E2.1 exists and μ_{k+1} is the measure induced on S_{k+1}.

PROOF Measurability of Γ_k means that μ_{k+1} defined by $\mu_{k+1}(Q)=\mu(\Gamma_k^{-1}Q)$ is a measure. But this is just E2.1.

The previous theorems depend entirely on the product topology assumed for the models. As a direct result, we need only consider the measurability of $\Phi_2=\Phi$ in trying to derive conditions applicable to whole models which will guarantee the existence of all induced measures in the models. The first result is an extension of the well known case when S_2 is a single operator:

THEOREM 3 Φ is measurable whenever S_2 is a separable discrete space.

PROOF S_2 must consist of a countable number of points $\{b_i\}$. Letting Q be a measurable set in S_3, $b_i^{-1}Q$ is a measurable set since all b_i are measurable. So,

$$\Phi^{-1}Q=\bigcup_i(b_i^{-1}Q,b_i)$$

is measurable.

This is true for completely arbitrary spaces S_1. Countability of S_1, on the other hand, is not sufficient to ensure measurability of Φ. For instance, let $S_1=\{1\}$, $S_3=\{0,1\}$ and $S_2(D)=\{b_x:x\varepsilon\underline{R}^*\}$ where the maps b_x are defined by

$$b_x(1)=\{1 \text{ if } x\varepsilon D, \text{ and } 0 \text{ otherwise}\}$$

and D is any subset of the extended real line $\underline{R}^*$. Metrics on S_1 and S_3 are trivial; on S_2 let

$$d_2(b_x,b_y)=|x-y|$$

so that S_2 is isomorphic and isometric to $\underline{R}^*$. If D is any non-measurable set then $\Phi^{-1}(\{1\})$ is non-measurable, even though each operator b_x in $S_2(D)$ is measurable.

The problem in this example is that $b_n \to b$ in S_2 does not imply that $b_n(x) \to b(x)$ in S_3. This difficulty need not arise...in fact, the following theorem shows that point-wise convergence of the operators in S_2 is sufficient for continuity, hence measurability of Φ. Necessary conditions for the measurability of Φ remain to be discovered, however.

THEOREM 4 If the maps in S_2 are continuous and convergence in S_2 is pointwise, then Φ is continuous, hence measurable.

PROOF Let $\{a_i\}$ be a subset of S_1, $a_i \to a$ and $\{b_j\}$ be a subset of S_2, $b_j \to b$. Then

$$d_3(\Phi(a_i,b_j),\Phi(a,b)) = d_3(b_j(a_i),b(a))$$
$$<= d_3(b_j(a_i),b_j(a)) + d_3(b_j(a),b(a))$$
$$< \varepsilon(a_i) + \varepsilon(b_j)$$

where the first term converges to zero by continuity of the b_i and the second by the point-wise convergence in S_2. This shows that Φ is a continuous map of $S_1 \times S_2 \to S_3$.

As an important example of a situation where S_2 is a space to which Theorem 4 applies, let S_1 and S_3 be metric spaces and S_2 the set of all continuous maps of S_1 into S_3 which have bounded ranges, meaning that the range of each $b \in S_2$ can be covered by a single ball of finite radius. S_2 becomes a metric space if the metric is:

$$d_2(b_1,b_2) = \sup_{a \in S_1} d_3(b_1(a),b_2(a))$$

Verifying the metric space axioms:

1. $d_2(b_1,b_2)>=0$ and $=0$ iff $b_1=b_2$

2. $d_2(b_1,b_2)=d_2(b_2,b_1)$

3. $d_2(b_1,b_2)+d_2(b_2,b_3)>=d_2(b_1,b_3)$

   which follows from

$$
\begin{aligned}
d_2(b_1,b_3) &= \sup_a d_3(b_1(a),b_3(a)) \\
&<= \sup_a (d_3(b_1a,b_2a)+d_3(b_2a,b_3a)) \\
&<= \sup_a d_3(b_1a,b_2a) + \sup_a d_3(b_2a,b_3a) \\
&= d_2(b_1,b_2)+d_2(b_2,b_3)
\end{aligned}
$$

As an easy consequence of these definitions we have:

THEOREM 5 If S_3 is complete, so is S_2.

As a complete metric space, S_2 supports a Borel σ-field and S_1, S_2, S_3 form a class of C/D models (when taken over all measures of weight 1 on the measurable product space S). Since convergence in S_2 is pointwise, Theorem 4 applies and a measure μ_3 is induced on S_3 by the model. Many other spaces have a suitable topology also. For instance, if S_1 and S_3 are Banach spaces then S_2 can be taken as the space of normed linear operators from S_1 into S_3, and Theorem 4 will apply.

The results given here are not inclusive, but they serve our immediate needs: all of the models used in the sequel will satisfy the conditions of Theorem 4. In the obvious cases, no mention will be made of this fact.

2.4 Two Examples of Classes of Models

Examples serve to clarify general concepts, so several will be presented using the ideas introduced in sections 2.1 and 2.2.

Detection of Gaussian Signals in Additive Gaussian Noise

Consider the problem of detecting a Gaussian signal obscured by additive Gaussian noise, but otherwise unaffected by transmission. A model for this situation follows: (the Synonym column references section 2.1 while the Space column references section 2.2)

Space	Synonym	Meaning and Definition
S_1	A	The set of source characters. Here, the set $\{1,0\}$ interpreted as {signal, no signal}.
S_2	E	Encoding operators. Here the signals are Gaussian, so, take E to be the set $\{e_h\}$ where h ranges over (abstract) Hilbert space, H. Then, for $a \epsilon A$, $e_h(a) \equiv a \cdot h$.
S_3	W	The channel waveforms. In this case, the Hilbert space H.
S_4	N	Noise operators. The noise is additive Gaussian, so take N to be the set $\{n_h : h \epsilon H\}$. For $w \epsilon W$, $n_h(w) \equiv w+h$.
S_5	R	Received waveforms. Again, just H, which recurs several times in this example because the transmission mapping is the identity operator.

Space	Synonym	Meaning and Definition
S_6	G	Detection operators. Frequently the detection operator is not stochastic; in this example assume that it is not and take $G=\{g\}$ for some fixed $g:R\to A$.
S_7	A	Estimates; the set $\{1,0\}$, interpreted as {signal present, no signal present}.

Specification of μ completes the model. Assuming independence (see E2.2) we have

$$\mu=\mu_1\mu_2\mu_4\mu_6$$

where

μ_1 is discrete, $\mu_1\{i\}=p_i$

μ_6 is degenerate, $\mu_6\{g\}=1$

μ_2 and μ_4 are independent Gaussian distributions.

The induced distributions μ_3, μ_5 and μ_7 are also of interest. μ_3 is not Gaussian because of the spike of mass p_o at the origin. Of course, μ_3^a, the distribution in S_3 conditioned by a value in S_1 is Gaussian. The same holds for μ_5: unconditioned it is not a normal distribution, but conditioned by a S_1, it is. Finally, $\mu_7(\hat{a}_i)=d_i$, the detection probability of $\hat{a}_i\varepsilon S_7$.

If H has finite dimension, then it is isomorphic and isometric to $\underline{R}^n$. Suppose then that μ_2 is the distribution which has autocovariance P and mean p while μ_4 has autocovariance Q and mean q. The densities are:

$$m_2(h)=1/\sqrt{((2\pi)^n|P|)}\ \exp(-(h-p,P^{-1}(h-p))/2)$$
$$m_4(h)=1/\sqrt{((2\pi)^n|Q|)}\ \exp(-(h-q,Q^{-1}(h-q))/2)$$

In order to derive μ_3, recall that $\chi_z(F)$ is the indicator function of the set F and is defined by

$$\chi_z(F)=\{1 \text{ if } z\varepsilon F, \ 0 \text{ otherwise}\}$$

Also, let

$$F_0=F-\{0\}$$

Then

$$\mu_3(F)=\int_{S_2} \mu_1(e^{-1}F)\mu_2(de)$$

$$=\int_{S_2} \mu_1(e^{-1}F_0)\mu_2(de) + \chi_0(F) \int_{S_2} \mu_1(e^{-1}\{0\})\mu_2(de)$$

$$=p_1 \int_{F_0} \mu_2(de) + p_0\chi_0(F)$$

or, in terms of densities:

$$m_3(x)=p_1m_2(x) + p_0\delta(x=0)$$

Similarly, μ_5 can be expressed as

$$\mu_5(F)=\int_{S_4} \mu_3(n^{-1}F)\mu_4(dn)$$

$$m_5(x)=\int_{S_4} \mu_3(x-n)\mu_4(n)dx$$

$$=m_3*m_4(x)$$

$$= p_1m_2*m_4(x) + p_0m_4(x)$$

Known Waveform in Additive Gaussian Noise Communications

This example is a slight modification of the previous one. S_1 becomes $\{0,1,2,\ldots,n\}$ interpreted as the set

{no character,character #1,...,character #n}

We take $\mu_1(i)=p_i$ as before, but the encoding operator is a fixed map $e:A\to W\subseteq H$, so that μ_2 is degenerate at the point e. This reflects the known waveform assumption. μ_4 and N are unchanged, but the induced distribution μ_5 is more complex, being conditionally Gaussian at each of the source characters. We may leave μ_6 unchanged although the detector, g, is a different operator. Finally, $\mu_7(i)=d_i$ as before.

2.5 Topologies on Classes of Models

When a class of models is to be manipulated, advantage can often be taken of structural properties of the class. For example, suppose that a performance criterion Π is defined on a class, C, of models. (By this we mean that Π is a bounded, real-valued function on the set C. For more detail, see section 2.7) If C has only the discrete topology, then selection of the model which maximizes Π can be done only by a straight search. On the other hand, if C is a normed linear space, a gradient search technique can be used. Since a class of models is really a set of measures, it is the [illegible] structures on sets of measures which must be investigated. We will look at two basic topologies for sets of probability measures. One well known method embeds them within a Banach space (the [illegible] space) while the other method, due to Kakutani [1], embeds them within abstract Hilbert space.

Embedding Within a Sup-norm Banach Space

It is well known that the probability measures can be given a metric derived from a Banach space of all finite measures. Here we look at one Banach space based on the sup-norm. Let M be the set of all signed, finite measures on a measurable space (X,S). M is a linear space over the reals, and defining

$$||\mu|| = \int_X d\mu_+ + \int_X d\mu_- = \int_X d|\mu|$$

we can easily verify that

1.	$||\mu|| >= 0$, and $=0$ iff $\mu \equiv 0$	true zero
2.	$||a \cdot \mu|| = |a| \cdot ||\mu||$	scalar multiplication
3.	$||\mu+\nu|| <= ||\mu|| + ||\nu||$	triangle inequality
	Proof: $d|\mu+\nu| <= d|\mu| + d|\nu|$	

so that M is a normed linear space. It is also a Banach space since it is complete:

THEOREM 6 M is complete

PROOF Let $\{\mu_i\}$ in M be a Cauchy sequence, $\mu_i \to \mu \varepsilon \overline{M}$ where $\overline{M}$ is the completion of M. We wish to show that μ is actually in M so that $\overline{M} = M$. But,

$$||\mu|| <= ||\mu - \mu_i|| + ||\mu_i||$$

and since $\{\mu_i\}$ is bounded (it is Cauchy), μ is finite.

It remains to show that μ is countably additive. Letting $\{B_i\}$ be a sequence of disjoint sets in S, set

$$Q_n = |\mu(\sum_{i=1}^{n} B_i) - \sum_{i=1}^{n} \mu B_i|$$

Now choose a subsequence of the μ_i, call it μ_k, so that

$||\mu_k - \mu||$ is non-increasing. (set $\mu_i \leftarrow -\mu_i$ if necessary) Choose a further subsequence μ_n from the sequence μ_k so that

$$||\mu_n - \mu|| < \varepsilon^n/2^n$$

for an arbitrary $\varepsilon > 0$. Then it is clear that

$$\sum_{i=1}^{n} |(\mu_n - \mu)B_i| <= \sum_{i=1}^{n} ||\mu_n - \mu|| = \varepsilon^n (1 - 2^{-n})$$

Since μ_n is countably additive,

$$Q_n = |\mu(\sum_{i=1}^{n} B_i) - \mu_n(\sum_{i=1}^{n} B_i) + \mu_n(\sum_{i=1}^{n} B_i) - \sum_{i=1}^{n} \mu B_i|$$

$$<= |(\mu - \mu_n)(\sum_{i=1}^{n} B_i)| + \sum_{i=1}^{n} |(\mu - \mu_n)B_i| < \varepsilon^n$$

Choosing $\varepsilon < 1$, $\lim_{n\to\infty} Q_n = 0$ and μ is countably additive.

The probability measures form a subset P within M. Since probability measures μ_p are distinguished by $d\mu_p >= 0$ and by $\int d\mu_p = 1$ or $||\mu_p|| = 1$, this subset is a small portion of the unit sphere in M. P satisfies the metric space axioms if $d(\mu,\nu) = ||\mu - \nu||$. P is bounded since

$$d(\mu,\nu) <= d(\mu,0) + d(0,\nu) = 2$$

Furthermore, P is complete since $\{\mu_i\}$ in P with $\mu_i \to \mu$ means that

$$||\mu|| <= ||\mu_i|| + ||\mu - \mu_i||$$

$$||\mu|| >= ||\mu_i|| - ||\mu - \mu_i||$$

which shows that $||\mu|| = 1$.

Best Available Copy

Lebesgue-Stieltjes Measures and the Space M′

When (X,S) is ($\underline{R}^n$,L^n), that is, Lebesgue measurable sets on n-dimensional Euclidean space, the space M is especially interesting since its elements can be represented explicitly as Lebesgue-Stieltjes measures. In applications, it is almost always this space which is used since explicit calculations can be made with relative ease. Often, (e.g., with Gaussian distributions) interest is centered on distributions which are continuously, or at least piecewise continuously, differentiable. These form a dense manifold M′ within M($\underline{R}^n$) which is not closed under the M-norm. Introducing a simple sup-norm on the derivatives, however, makes M′ itself into a Banach space. The set P′ derived from P in a like manner is a subset of M′ of course. While P is bounded, P′ is not, although it is a metric space. However, P′ is complete:

LEMMA 1 If $p_n' \to p'$ in P′ then $p_n \to p$ in P where $p_n = \int p_n'$, $p = \int p'$

PROOF The integral is continuous.

THEOREM 7 P′ is complete.

PROOF Let $\{p_i'\}$ be a Cauchy sequence in P′, that is, $p_i' \to p' \epsilon \overline{P}'$, the completion of P′. p′ is piecewise continuous because of the sup-metric in P′ . By lemma 1, $\int p=1$ so that $p' \epsilon P'$, or, $P' = \overline{P}'$, as required.

Embedding Within Hilbert Space (Kakutani [1])

The set P defined on the previous page can also be embedded within a Hilbert space in order to form a metric space P_H. Consider first the case $\mu \sim \nu$ and define

$$\rho(\mu,\nu)= \int_X \sqrt{(d\mu/d\nu)}\, d\nu$$

From Schwarz's inequality we can show that

1. $0<\rho(\mu,\nu)<=1$

2. $\rho(\mu,\nu)=1$ iff $\mu=\nu$

3. $\rho(\mu,\nu)=\rho(\nu,\mu)$

Proof: $\rho(\mu,\nu)= \int_X \sqrt{(d\mu/d\nu)}\, d\nu= \int_X (1/\sqrt{(d\nu/d\mu)})(d\nu/d\mu)\, d\mu$

$$= \int_X \sqrt{(d\nu/d\mu)}\, d\mu=\rho(\nu,\mu)$$

Now we can obtain a quasi-metric (i.e., a metric without the triangle inequality) by

$$\sigma(\mu,\nu)=-\mathrm{Ln}\ \rho(\mu,\nu)= -\mathrm{Ln} \int_X \sqrt{(d\mu/d\nu)}\, d\nu$$

since we have

1. $0<=\sigma(\mu,\nu)<\infty$

2. $\sigma(\mu,\nu)=0$ iff $\mu=\nu$

3. $\sigma(\mu,\nu)=\sigma(\nu,\mu)$

only the triangle inequality fails at times to hold. It is the logarithm in the definition of σ which distorts the "distance" surface. For instance, let $X=\{x_1,x_2\}$ and suppose that:

	x_1	x_2
μ	3/4	1/4
ν	1/2	1/2
ω	1/4	3/4

Direct calculation shows that

$$\rho(\mu,\nu)=\rho(\nu,\omega)= 0.965$$

$$\rho(\mu,\omega)= \qquad 0.816$$

so that

$$\sigma(\mu,\nu) + \sigma(\nu,\omega) = 0.092 < 0.14 = \sigma(\mu,\omega)$$

which is a clear failure of the triangle inequality.

Now let $\mu,\mu' \epsilon P_H$ be arbitrary. We do not assume that $\mu \sim \mu'$. Choose a third measure $\nu \epsilon P_H$ such that $\mu < \nu, \mu' < \nu$. $(\mu+\mu')/2$ is one such element, but there may be many others. Now define two elements of $L^2(X,S,\nu)$ by

$$\Psi(\omega)=\sqrt{(d\mu/d\nu)}(\omega) \qquad \Psi'(\omega)=\sqrt{(d\mu'/d\nu)}(\omega)$$

Both Ψ and Ψ' are on the unit sphere in L^2, and, when $\mu \sim \mu'$ it is clear that

$$\rho(\mu,\mu')= \int_X \sqrt{(d\mu/d\mu')}d\mu'= \int_X \sqrt{(d\mu/d\nu)}\sqrt{(d\mu'/d\nu)}d\nu= (\Psi,\Psi')$$

where (Ψ,Ψ') is the inner product of Ψ and Ψ' in L^2.

This relationship provides a natural extension of $\rho(\mu,\mu')$ to those cases where it is not the case that $\mu \sim \mu'$. It is also clear that $\rho(\mu,\mu')$ so defined is independent of ν so long as $\mu < \nu$ and $\mu' < \nu$. Also, only when $\mu \perp \mu'$ does $\rho(\mu,\mu')=0$. Finally, since

$$||\Psi-\Psi'||^2=||\Psi||^2+||\Psi'||^2-2(\Psi,\Psi')=2(1-\rho(\mu,\mu'))$$

we see that P_H can be made into a metric space by defining

$$d(\mu,\mu')=||\Psi-\Psi'||=\sqrt{(2(1-\rho(\mu,\mu')))}$$

Since this metric is independent of the choice of ν, P_H has been isometrically embedded into abstract Hilbert space.

The quasi-metric σ can also be useful, especially for Gaussian measures, since it induces the same topology on P_H as does $d(\cdot,\cdot)$.

This follows since there exist two constants k_1 and k_2 such that

$$k_1\sigma(\mu,\mu')<=d^2(\mu,\mu')<=k_2\sigma(\mu,\mu')$$

for either $\sigma(\mu,\mu')$ or $d(\mu,\mu')$ sufficiently small.

2.6 Application to Examples

It has only been possible to apply the metrics of the previous section to concrete numerical examples in the case of independent models, where each marginal measure can be treated independently. The whole model can then be treated as one point in the appropriate product metric space, although we do not explicitly do that here, but are content to develop the metrics for the individual marginal measures. The forms of the P-space, P′-space and P_H-space (Hilbert) metrics are developed for countable spaces and for one-dimensional Gaussian distributions.

Sequence Spaces

For distributions on denumerable spaces, such as μ_1 on S_1 in the examples of section 2.4, we define $p_i=\mu\{i\}$. Then it is clear that: (define $q_i=\nu\{i\}$)

Denumerable P-space Sup-metric

$$||\mu-\nu||=\max_i|p_i-q_i|$$

Denumerable P_H-space Hilbert Metric

$$d(\mu,\nu)=[2-2\rho(\mu,\nu)]^{1/2}=[2-2\sum_i\sqrt{p_i}\sqrt{q_i}\ ]^{1/2}=[\sum_i(\sqrt{p_i}-\sqrt{q_i}\ )^2]^{1/2}$$

Gaussian Case

Gaussian distributions on $\underline{R}^n$ are also interesting. Let F and G be Gaussian distributions, means m_F and m_G, variances σ_F and σ_G, and with densities dF and dG. We define Q_1 and Q_2 to be the roots of the quadratic equation in the variable Q:

$$\text{Ln } dF(Q)=\text{Ln } dG(Q) \tag{E2.6}$$

Since, for $K\epsilon\{F,G\}$,

$$dK(Q)=(\sigma_K\sqrt{2\pi}\ )^{-1}\ \exp[-(Q-m_K)^2/2\sigma_K^2]$$

we have this form of E2.6:

$$\text{Ln } \sigma_F + (Q-m_F)^2/2\sigma_F^2 = \text{Ln } \sigma_G + (Q-m_G)^2/2\sigma_G^2$$

If $\sigma_F=\sigma_G$ then $Q_1=(m_F+m_G)/2$, $Q_2=\infty$ (Figures 2.4 and 2.5)

Gaussian Measures in P-space over $\underline{R}^1$

In order to obtain an expression for the metric in P, we want to form $|F-G|(\underline{R}^1)$. Looking at the nine cases,

$$(\sigma_F <,=,> \sigma_G)x(m_F <,=,> m_G)$$

we see that $\sigma_F<\sigma_G$ means that dF>dG between Q_1 and Q_2 no matter what values m_F and m_G have and that dF<dG between Q_1 and Q_2 whenever $\sigma_F>\sigma_G$. When $\sigma_F=\sigma_G$, dF>dG up to Q_1 if $m_F<m_G$ and dF<dG up to Q_1 if $m_F>m_G$. If $m_F=m_G$ then dF=dG everywhere.

Consider case #1: $\sigma_F>\sigma_G$

$$d(F,G)=|F-G|(\underline{R}^1)=\int d|F-G|$$

$$=\int_{-\infty}^{Q_1}(dF-dG) + \int_{Q_1}^{Q_2}(dG-dF) + \int_{Q_2}^{+\infty}(dF-dG)$$

Defining

$$\text{erf}(x)=\int_x^{+\infty}\sqrt{2\pi}^{\,-1}\exp(-\xi^2/2)d\xi$$

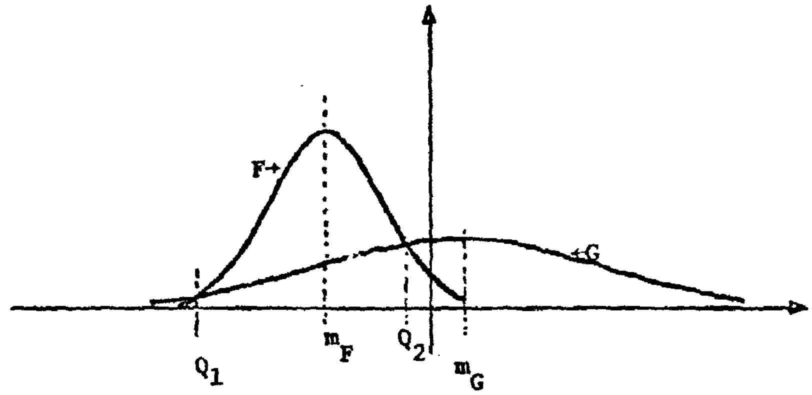


Figure 2.4

Roots of E2.6 when $\sigma_F > \sigma_G$, $m_F < m_G$

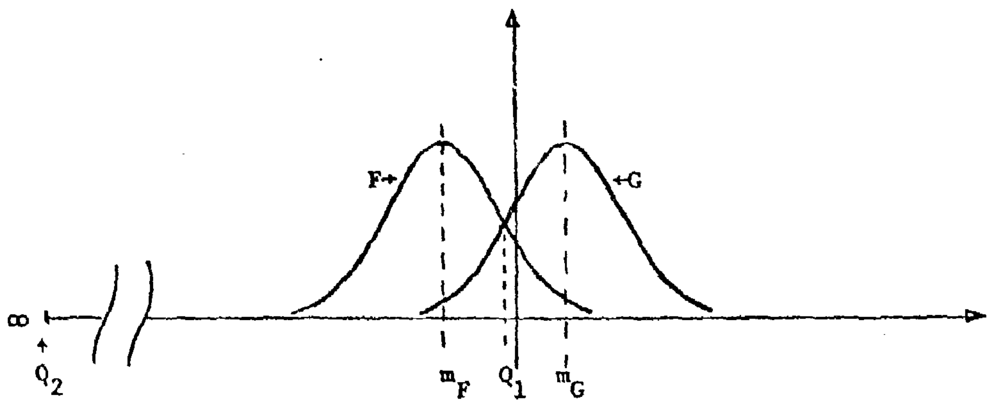


Figure 2.5

Roots of E2.6 when $\sigma_F = \sigma_G$, identifying $+\infty$ and $-\infty$,
i.e., using the one point compactification of the real line.

we see that

$$\int_{-\infty}^{y} \sqrt{2\pi}^{\,-1} \exp(-\xi^2/2)d\xi = 1-\mathrm{erf}(y)$$

$$\int_{x}^{y} \sqrt{2\pi}^{\,-1} \exp(-\xi^2/2)d\xi = \mathrm{erf}(x)-\mathrm{erf}(y)$$

so that

$$\begin{aligned} d(F,G) = & -2\mathrm{erf}[(Q_1-m_F)/\sigma_F] + 2\mathrm{erf}[(Q_2-m_F)/\sigma_F] \\ & +2\mathrm{erf}[(Q_1-m_G)/\sigma_G] - 2\mathrm{erf}[(Q_2-m_G)/\sigma_G] \end{aligned}$$

and since $\sigma_F<\sigma_G$ changes the signs, case #2 gives the negative of the above equation so that, for $\sigma_F \neq \sigma_G$:

$$\begin{aligned} d(F,G) = 2\,| & \mathrm{erf}[(Q_1-m_F)/\sigma_F] - \mathrm{erf}[(Q_1-m_G)/\sigma_G] \\ & +\mathrm{erf}[(Q_2-m_G)/\sigma_G] - \mathrm{erf}[(Q_2-m_F)/\sigma_F]\,| \end{aligned}$$

while if $\sigma_F=\sigma_G$ then

$$d(F,G)=2\,|\mathrm{erf}[\, m_G \,/\, 2\sigma \,]-\mathrm{erf}[\, m_F \,/\, 2\sigma \,]\,|$$

These two equations define $d(\cdot,\cdot)$ for any Gaussian measures in $P(\underline{R}^1)$.

Gaussian Measures in P'-space over $\underline{R}^1$

While the P-space distance is relatively messy, the P′ distance between Gaussian distributions has a particularly simple form. Letting f and g be the frequency functions of the distributions F and G we have

$$\begin{aligned} d(f,g) &= \sup_x |f(x)-g(x)| \\ &= \max\{f(m_F)-g(m_F),\, g(m_G)-f(m_G)\} \\ &= \{f(m_F)-g(m_F) \quad \text{when } \sigma_F \leq \sigma_G \text{ and} \\ &\quad\; g(m_G)-f(m_G) \quad \text{when } \sigma_F \geq \sigma_G\} \end{aligned}$$

Gaussian Measures in P_H-space over $\underline{R}^1$

Finally, we evaluate the metric derived from embedding the set P within Hilbert space. Computing $\rho(\cdot,\cdot)$:

$$\rho(F,G)=\int\sqrt{dF/dG}\,dG$$

$$=(1/\ 2\pi\sigma_F\sigma_G\ )\int\exp[-[(x-m_F)/2\sigma_F]^2-[(x-m_G)/2\sigma_G]^2]dx$$

completing squares in the integrand,

$$=(\sigma_F\sigma_G\sqrt{2}\ /\sqrt{\sigma_F^2+\sigma_G^2}\ )\exp[-\sigma_F^2\sigma_G^2(m_F-m_G)^2/(\sigma_F^2+\sigma_G^2)^2]$$

and substituting we obtain:

$$d(F,G)=(2[1-\rho(F,G)])^{1/2}$$

$$=\sqrt{2}\ [1-(\sigma_F\sigma_G\sqrt{2}\ /\sqrt{\sigma_F^2+\sigma_G^2}\ )\exp[-\sigma_F^2\sigma_G^2(m_F-m_G)^2/(\sigma_F^2+\sigma_G^2)^2]]^{1/2}$$

which shows just how complicated conceptually simple results can become. As a next step, these formulae should be extended to multi-dimensional Gaussian distributions, but we will not do it here.

2.7 Performance Criteria and Linear Risk

No sooner are we given a class C of models then we want to choose one of the class to analyze or perhaps to build. The easiest way to accomplish this is to define a real valued function Π on C. The set-valued right inverse Π^{-1} then images the total ordering < on $\underline{R}^1$ into a total ordering << defined on C. Either the lub or glb of << should exist in C so that this element can be the one chosen. Since changing $\Pi\leftrightarrow-\Pi$ switches the lub and glb of the induced ordering <<, we will be interested in showing that Π is bounded either above or below, but not necessarily both.

We will only discuss a single class of performance criteria, but they are useful for simple detection problems since they take $S_{2L+1}=S_1$ rather than some combination of the intermediate spaces in which case the model would be set up for estimation problems or mixed detection/estimation problems. We arrive at our class of criteria by assuming a linear relationship between the various detection alternatives and the total benefit produced by the model.

Let $C(a,b)$ be the benefit derived when $a\varepsilon S_1$ is transmitted and $b\varepsilon S_{2L+1}=S_1$ is detected. The expected benefit is just

$$\Pi_C(\mu)=\int C(a,b)\mu^a_{2L+1}(db)\mu_1(da)$$

where the integration is over $S_1 x S_{2L+1}$ and the integral exists whenever $C(a,b)$ is measurable. When $C(a,b)$ is bounded, Π_C is too. This Baysian performance criterion can also be written in an expanded form as

$$\Pi_C(\mu)=\int C(a,b)\chi_a(\underline{t}^{-1}db)\mu(dad\underline{t})$$

Π_C should be insensitive to changes in μ since μ can never be known exactly. At the least this should mean continuity of Π_C with respect to some topology on the space of which μ is a member. One of the merits of a linear risk (or benefit) performance criterion is precisely this continuity in both of the spaces P and P′:

THEOREM 8 If $|C(a,b)|$ is bounded by $\bar{c}$ then Π_C is continuous in P.

PROOF
$$\begin{aligned}|\Pi(\mu)-\Pi(\nu) &=|\int C(a,b)\chi_a(\underline{t}^{-1}db)(\mu-\nu)(dad\underline{t})\\ &<=\bar{c}\int\chi_a(\underline{t}^{-1}db)|\mu-\nu|(dad\underline{t})\\ &=\bar{c}\int|\mu-\nu|(dad\underline{t})\\ &=\bar{c}d_P(\mu,\nu)\end{aligned}$$

where $d_P(\cdot,\cdot)$ is the metric in P.

COROLLARY 1 If S_1 is countable then $\mu^a_{2L+1}(b)$ is continuous in P.

COROLLARY 2 If C(a,b) is bounded then Π_C is continuous in P'.

PROOF Let $p'_n \to p'$ in P'. We wish to show that $\Pi_C(p_n)\to\Pi_C(p)$ where $p_n = \int p'_n$ is in P. But since $p'_n \to p'$ in P' implies that $p_n \to p$ in P (Lemma 1, section 2.5), this follows immediately.

2.8 Expected Error and Singularity

An important special case of linear benefit arises when $C(a,b)=d_1(a,b)$ where $d_1(\cdot,\cdot)$ is the metric in the spaces S_1 and S_{2L+1}. The resulting performance criterion, Π_{d_1}, or simply Π_d, is the expected error. If $\Pi_d(\mu)$ is zero, the model

$$M=(S_1,S_2,\ldots,S_{2L+1},\mu)$$

is said to be singular. If $\mu_1(a)>0$ then a is called a naturally occurring source character; it is clearly no restriction to consider only naturally occurring source characters. The following theorem and corollary are basic to an understanding of model singularity.

THEOREM 9 If $S_1=S_{2L+1}$ are countable with a discrete metric then M is singular iff $\mu^r_{2L+1}(r)=1$ for every naturally occurring $r\varepsilon S_1$.

PROOF To establish necessity, assume that M is singular. Then

$$0=\Pi_d(\mu)=\int d(a,b)\mu^a_{2L+1}(db)\mu_1(da)$$

$$=\sum_{i,j} d(a_i,b_j)\mu^{a_i}_{2L+1}(b_j)\mu_1(a_i)$$

and since each term in the sum is positive, each term must be zero separately:

$$d(a_i,b_j)\mu^{a_i}_{2L+1}(b_j)\mu_1(a_i)=0 \quad \text{for all } i,j$$

from which the necessity follows, and, so does the sufficiency.

COROLLARY 1 If $S_1=S_{2L+1}$ are countable with a discrete metric then M is singular iff $\mu^r_{2L+1}\perp\mu^s_{2L+1}$ for every naturally occurring pair $r\neq s$ in S_1.

This corollary expresses the well known, but usually imprecisely stated, fact that model singularity is equivalent to measure orthogonality. An obvious extension of this theorem is provided by:

THEOREM 10 If M factors then $\mu^r_{2L+1}\perp\mu^s_{2L+1}$ for a naturally occurring pair $r\neq s$ iff $\mu^r_{2k+1}\perp\mu^s_{2k+1}$ for all $k<L$.

PROOF We will establish necessity for $k=L-1$ by showing the contrapositive. Letting $Q^r_k=\text{supp}(\mu^r_{2k+1})^{\#}$, this means assuming either $\mu^r_{2k+1}(Q^s_k)>0$ or $\mu^s_{2k+1}(Q^r_k)>0$ and showing that the same holds for either μ^r_{2L+1} or μ^s_{2L+1}, respectively. First note that

$$\mu^r_{2L+1}(R)=\int_{S^*_1}\chi_r(\underline{t}^{-1}R)\mu^*_{2L}(d\underline{t})$$

\# This and subsequent proofs have been simplified by suppressing references to "some" support and talking instead about a (reads like the) support of a measure. No error has been introduced by doing this, while the arguments have become easier to follow.

or, since M factors and 2L-1=2k+1

$$\mu^r_{2L+1}(R) = \int_{Q^r_k \times S_{2L}} \chi_a(t^{-1}_{2L}R)\mu_{2L}(dt_{2L})\mu^r_{2L-1}(da)$$

and a similar expression holds for $\mu^s_{2L+1}(R)$.

Now assume that $\mu^r_{2k+1}(Q^s_k)>0$ and look at:

$$\mu^r_{2L+1}(Q^s_L) = \int_{Q^r_k \times S_{2L}} \chi_a(t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})\mu^r_{2L-1}(da)$$

$$= \int_{S_{2L}} \mu^r_{2k+1}(t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

$$>= \int_{S_{2L}} \mu^r_{2k+1}(Q^s_k \cap t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

which by the lemma below, is greater than zero. This establishes necessity and hence the theorem, since the sufficiency is immediate from the definitions.

LEMMA 1 If $\mu^r_{2k+1}(Q^s_k)>0$ then

$$0<\int_{S_{2L}}\mu^r_{2k+1}(Q^s_k \cap t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

PROOF

$$1 = \mu^s_{2L+1}(Q^s_L) = \int_{S_{2L}}\mu^s_{2k+1}(t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

$$= \int_{S_{2L}}\mu^s_{2k+1}(Q^s_k \cap t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

hence $t^{-1}_{2L}Q^s_L \subset Q^s_k$ for all $t_{2L}\in\mathrm{supp}(\mu_{2L})$. But, stronger than that, this says that Q^s_k is covered by the inverse maps of Q^s_L, or, precisely,

$$\bigcup_{t_{2L}\in\mathrm{supp}(\mu_{2L})} t^{-1}_{2L}(Q^s_L) \supset Q^s_k$$

But that means that if $\mu^r_{2k+1}(Q^s_k)>0$ then likewise

$$0< \int_{S_{2L}}\mu^r_{2k+1}(Q^s_k \cap t^{-1}_{2L}Q^s_L)\mu_{2L}(dt_{2L})$$

This theorem says that the conditional measures must start out orthogonal and stay orthogonal if the model is singular, and vice-versa. But, this holds only for factorable models, where the marginal measures are all independent. A non-factorable model can be rigged, as in the example of the next section, which is still singular even though some of the intermediate conditional measures are not. The converse counter-example which corresponds to this example is easily constructed, so it will not be explicitly given.

2.9 Singularity in a Non-factorable Model

Our example will illustrate the necessity of the factorability assumption in Theorem 10, section 2.8. Take as the model

$$M=(S_1,S_2,S_3,S_4,S_5,\alpha\cdot\nu)$$

where α is defined on S_1 and ν is defined on $S_2 x S_4$, and

$$S_1=S_3=S_5=\{0,1\}$$

$$S_2=S_4=\{0,1\}x\{0,1\}$$

The measure α is given by

$x\varepsilon S_1$	$\alpha(x)$
1	1/2
0	1/2

S_2 and S_4 are both spaces of operators, and in order to give ν on $S_2 x S_4$, we will represent pairs of operators $(b,d)\varepsilon S_2 x S_4$ by enumeration of the operator values on all of the points in $S_2 x S_4$. To define one pair (b,d) we only need to give four values since S_1, the domain of $b\varepsilon S_2$, and S_3, the domain of $d\varepsilon S_4$, each consist of only

two points. All possible operator pairs (b,d) can thus be represented by all possible 4-tuples with entries of zero and one. So, letting each $(b,d)\varepsilon S_2 x S_4$ be represented by the 4-tuple of values (b(0),b(1),d(0),d(1)), we define ν by this table of values:

$x\varepsilon S_2xS_4=(b(0),b(1),d(0),d(1))$				$\nu(x)$
0	0	0	0	
0	0	0	1	
0	0	1	0	
0	0	1	1	
0	1	0	0	
0	1	0	1	1/2
0	1	1	0	
0	1	1	1	
1	0	0	0	
1	0	0	1	
1	0	1	0	1/2
1	0	1	1	
1	1	0	0	
1	1	0	1	
1	1	1	0	
1	1	1	1	

The marginal distributions can then be calculated:

$x\varepsilon S_2$	$\beta(x)=\int_{S_4}\nu(x,d\omega)$
0 0	0
0 1	1/2
1 0	1/2
1 1	0

$x\varepsilon S_4$	$\delta(x)=\int_{S_2}\nu(d\omega,x)$
0 0	0
0 1	1/2
1 0	1/2
1 1	0

and we can calculate the conditional induced measures as well. On the intermediate space, S_3, we have:

i j	$\mu_3^i(j)=\sum_{k=0}^{3}\chi_i(b_k^{-1}(j))\beta(b_k)$
0 0	1/2
0 1	1/2
1 0	1/2
1 1	1/2

This shows clearly that $\mu_3^0 \sim \mu_3^1$, in fact, they are equal. However, on the last space, S_5, the conditional induced measures are:

i j	$\mu_5^i(j) = \sum_{k,m=0}^{3} \chi_i(b_k^{-1} \circ d_m^{-1}(j))\nu(db_k, dd_m)$
0 0	1
0 1	0
1 0	0
1 1	1

clearly showing that $\mu_5^0 \perp \mu_5^1$.

This, then, is a non-factorable 2 stage model. The marginal measures after the first stage are equivalent, while the marginal measures after both stages are orthogonal, making the model singular. This happens because the two stages are not independent; the model does not factor. As a result, the second stage can be arranged (and is in this example) so as to undo the random selection introduced by the first stage. Fortunately, this kind of dependence, correlation of 1, is not even a plausible representation of reality, and so would never be used in an actual C/D model.

3. LINEAR TRANSMISSION AND ADDITIVE NOISE

In chapter 2 the outlines of a general theory of communication and detection models were established. In this chapter that theory will be specialized in order to obtain certain results for models of additive noise and linear transmission. The assumption of additive noise is widely made because of the mathematical tractability which it provides...and, once additive noise is assumed, the assumption of linear transmission often follows.

A special "multiplicative" stage is also defined and analyzed. Although this type of stage is rather simple, it is included here for reference in chapter 4, where these three kinds of stages, multiplicative, additive, and linear, will be used in the synthesis of sonar models. The theorems in this chapter will then make it possible to expose and understand a certain kind of (additive) singularity that can arise in these models.

3.1 Additive Stages

When the effect of a stage of a C/D model is to add two linear subspaces together, as in stage 2 of the examples in section 2.4, a very important kind of singularity of the model can occur. Following Figure 3.1, suppose S and N are subspaces of a linear space, $S \cap N$ the subspace they have in common and S+N their sum. If μ_s is supported by $S-\{0\}$, (i.e., by S exclusive of the origin) $\mu_0(0)=1$ and $\text{supp}(\mu_n)\subset N$, we can consider two measures, μ_{S+N} and μ_{0+N} on the sum space S+N which are induced by addition of the

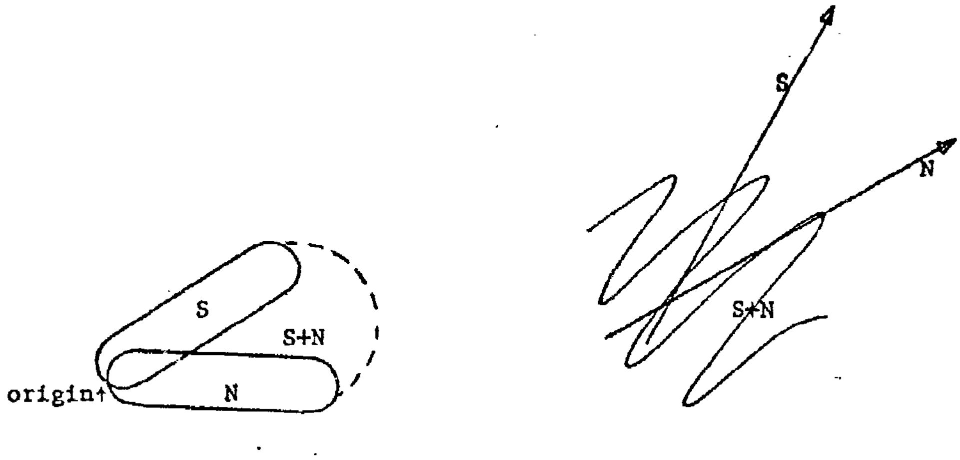


Figure 3.1

Possible Singularity in Additive Stages

two spaces. We see that $\mu_{0+N}=\mu_N$ and that μ_{S+N} is supported by that portion of S+N which is above N <u>if</u> $S\cap N=0$. The result is that $\mu_{S+N} \perp \mu_N$, and the reason is that $\mathrm{supp}(\mu_{S+N})$ has a linear projection outside of the subspace spanned by $\mathrm{supp}(\mu_N)$. This last phrase will turn up again in a stronger form when we consider Gaussian processes a little later in this section.

We say that the k-th stage, $(S_{2k-1}, S_{2k}, S_{2k+1})$ of a C/D model is additive if:

1. S_{2k+1} is a Banach space.
2. S_{2k-1} is a subspace of S_{2k+1}
3. $S_{2k}=\{b_h : h\varepsilon Q\}$ where Q is a subspace of S_{2k+1} and $b_h : S_{2k-1} \to S_{2k-1}+h \subset S_{2k+1}$

For simplicity of nomenclature, we will use S_{2k} to mean both the space of operators and the subspace Q of S_{2k+1} since the map $\Psi(b_h)=h$ is an isomorphism. Giving S_{2k} the norm $||b_h||=||h||$ makes Ψ an isometry as well.

As defined here, additive stages have an important property: when $S_{2k} \cap S_{2k-1}=0$, that is, when they have only the origin in common, singularity is preserved by the k-th stage. In an abuse of language that is unlikely to cause confusion, a stage which preserves singularity is itself said to be singular. If we let M_k be the first k stages of a C/D model M, we can state this as a theorem:

THEOREM 1 If M_{k-1} is singular and stage k is additive with $S_{2k} \cap S_{2k-1}=0$, then M_k is singular, i.e., stage k is singular.

PROOF Singularity of M_{k-1} implies the existence of disjoint supports for the conditional measures: $Q^s_{k-1} \cap Q^r_{k-1}=\{\}$, for all $s\neq r$. Since we also have $S_{2k} \cap S_{2k-1}=0$, we have

$$\bigcup_{b\varepsilon S_{2k}} b(Q^s_{k-1}) \cap \bigcup_{b\varepsilon S_{2k}} b(Q^r_{k-1})=\{\} \quad \text{for all } s\neq r$$

But,

$$\bigcup_{b \in S_{2k}} b(Q^s_{k-1}) \supset Q^s_k$$

so

$$Q^s_k \cap Q^r_k = \{\}, \text{ for all } s \neq r$$

and M_k is singular. (remember that Q^r_k is defined to be the support of μ^r_{2k+1}, see theorem 2.10)

This theorem is useful whenever additive noise is used in a model, as it is in the examples of section 2.4 and as it is in the sonar models of sections 4.4 and 4.5. Having these sufficient conditions for an additive stage to singular whets the appetite and motivates a search for necessary conditions. As a step in that direction:

THEOREM 2 If M_{k-1} is singular, stage k is such that $\text{supp}(\mu_{2k}) \supset S_{2k-1}$ and μ_{2k} is independent of μ_i for all $i<2k$ then M_k is non-singular.

PROOF

$$\mu^r_{2k+1}(Q^s_k) = \int_{Q^r_{k-1} \times S_{2k}} \chi_a(t^{-1}_{2k} Q^s_k)\mu_{2k}(dt_{2k})\mu^r_{2k-1}(da)$$

and by the assumed property of stage k,

$$\text{supp}(\mu_{2k}) \cap \text{supp}(\mu_{2k-1}) \neq \{\}$$

so that $t^{-1}_{2k} Q^s_k \cap Q^r_{k-1} \neq \{\}$ for all t in some set T of positive μ_{2k} measure in S_{2k}. Hence $\mu^r_{2k+1}(Q^s_k) > 0$ and M_k is non-singular.

Notice that this proof really says nothing about the singularity of M_{k-1}, so the theorem holds even if M_{k-1} is non-singular.

Gaussian Measures

When Gaussian measures are assumed in anything, stronger results can usually be expected. In the Gaussian detection models, in fact, M_k must either be singular or the conditional measures must be equivalent:

THEOREM 3 If ν and μ are Gaussian, either $\mu \perp \nu$ or $\mu \sim \nu$.

A proof in the general case has been given by J. Feldman [1]. For further discussion of related work, see section 1.3.

Various ways of telling whether $\mu \perp \nu$ or $\mu \sim \nu$ have been discovered. They fall, generally, into two groups: 1) general, universally applicable and non-constructive, hence useless in practical analysis. Feldman's original proofs are of this nature; 2) constructive, but applicable only to special kinds of normal distributions, such as Markov processes, or, stationary processes at least one of which has a rational spectral density. What is perhaps the most useful of the universal results was obtained by Kallianpur and Oodaira [1]:

THEOREM 4 $P \sim Q$ iff

1/ $m(\cdot) \in H(\Gamma_Q)$

2/ Γ_P has a representation $\Gamma_P(s,t) = \Sigma \mu_k e_k(s) e_k(t)$ where $\{e_k\}$ is a c.o.n. in $H(\Gamma_Q)$ and $\Sigma(1-\mu_k)^2 < \infty$ and $\mu_k >= c > 0$ for all k

In this theorem, $H(\Gamma)$ is the reproducing kernel Hilbert space (Aronsajn [1][2]) with kernel Γ and P,Q are the distributions for the normal processes with correlations Γ_P, Γ_Q and mean functions $m(\cdot)$ and 0 respectively.

As an immediate application to additive stages, we have:

COROLLARY 1 If $P \perp Q$ because $m(\cdot) \notin H(\Gamma_Q)$ then $P+Q \perp Q$

This can be applied to example number 1, section 2.4, to show that the second stage, which is additive, is singular if $\mu_3^1 \perp \mu_4$ because the mean of μ_3^1 is not in $H(\Gamma_{\mu_4})$. As an application of the second condition of Theorem 3 we have

COROLLARY 2 If $P \perp Q$ because $\mu_k \to 0$ or some $\mu_k = 0$ then $P \perp P+Q$

This is the situation that occurs when the signal, Q, occupies some dimensions ("bandwidth") that the noise does not. While this can not happen in a practical sense, it can plague model builders. More about this in Chapter 4. Finally, as a converse to Corollary 2:

COROLLARY 3 If $P \perp Q$ because $\mu_k \to 0$ or some $\mu_k = 0$ then $P+Q \sim Q$.

This is the normal situation: the noise, Q is wider-band than the signal, P, so the additive noise stage is non-singular.

Thinking in terms of linear spaces, Theorem 4 says to first look at the linear space most closely tied to the process (with distribution) Q. This is just $H(\Gamma_Q)$, the RKHS (Reproducing Kernel

Hilbert Space) with kernel Γ_Q. $H(\Gamma_Q)$ is a natural space for Q because A) it supports Q and B) distances reflect the concentration of Q along various axes. Now, given another Gaussian process P with mean $m(\cdot)$, in order to see if $P \perp Q$ we have to do several things. The first is to check that $m(\cdot)\epsilon H(\Gamma_Q)$. If it is not, then P is lifted up out of $H(\Gamma_Q)$, that is, has a linear projection outside of the support of Q, so, $P \perp Q$. If $m(\cdot)\epsilon H(\Gamma_Q)$ then we still have to check further. To begin with, Γ_P may not be representable in $H(\Gamma_Q)xH(\Gamma_Q)$, that is, Γ_P may have a linear projection outside of $H(\Gamma_Q)^2$. The support of P will then too, so that $P \perp Q$. Secondly, some μ_k may be zero. That is, $H(\Gamma_Q)$ may have linear dimensions not needed to represent Γ_P, in which case supp(Q) will have linear dimensions outside of supp(P), so that again $P \perp Q$. Finally, we are asked to look at the distribution of "energy" into the different "eigenfrequencies" of the two processes. Unless the two put almost the same energy on all but a finite number of dimensions, i.e., unless $\Sigma(1-\mu_k)^2<\infty$, then $P \perp Q$. While this last requirement has the only probabilistic flavor of all the requirements for $P\sim Q$, even it is closely related to the concept of projections ouside of a given linear space. What it says is that supp(P) and supp(Q) may not sneak linear projections outside of each other through divergent behavior at infinity if $P\sim Q$ is to hold.

3.2 Linear Stages

When S_{2k-1} and S_{2k+1} are Banach spaces and all of the elements in S_{2k} are linear operators (additive and continuous) from S_{2k-1} into S_{2k+1} then stage k is linear. As one expects, linear stages are also capable of preserving singularity, or of being singular to use the verbal shorthand introduced in section 3.1. The simplest case arises when μ_{2k} is degenerate at t, in which case the effect of the stage depends entirely upon the nullspace of t, call it T_0. Letting $\overline{Q_k^r}$ be the subspace spanned by Q_k^r, we have:

THEOREM 5 If S_{2k+1} is a Hilbert space, M_{k-1} is singular and $\overline{Q}_{k-1}^{r} \perp T_0$ for all but one $r \epsilon S_1$ then M_k is singular.

PROOF $t(\overline{Q}_{k-1}^{r}) \supset Q_k^r$ and $t(\overline{Q}_{k-1}^{s} \cap t(\overline{Q}_{k-1}^{r}) = 0$ since t is 1:1 on $S_{2k-1}-T_0$.

Extension of this result to Banach spaces requires the assumption that projections P_0 onto T_0 and $P_1=I-P_0$ exist, since the existence of a projection onto an arbitrary subspace of a Banach is not guaranteed. If P_1 does exist, we have:

THEOREM 6 If M_{k-1} is singular and $P_1Q_{k-1}^s=Q_{k-1}^s$ for all but one $s\epsilon S_1$ then M_k is singular.

Either of these theorems has an obvious extension to multiple operators via the simple expedient of defining:

$T_0=\{\text{union of all nullspaces of operators in } \mathrm{supp}(\mu_{2k})\}$

Under this definition, T_0 may not be unique, so the extended theorem must be phrased as: M_k is singular if M_{k-1} is singular and there exists at least one T_0 such that $P_1 Q^s_{k-1} = Q^s_{k-1}$ for all but one $s \epsilon S_1$.

3.3 Multiplicative Stages

The k-th stage of a model is said to be source multiplicative if:

1. S_{2k+1} is a Banach space
2. S_{2k} is isomorphic to S_{2k+1} in the usual way, see section 3.1, and $e_h \epsilon S_{2k}$ maps $a \epsilon S_{2k-1}$ into $a \cdot h \epsilon S_{2k+1}$.
3. S_{2k-1} is the field from which S_{2k+1} is generated, either $\underline{C}^1$ or $\underline{R}^1$.

The following almost trivial theorem on singularity of a source multiplicative stage, finds application to the first, or encoding stages of models in sections 4.4 and 4.5.

THEOREM 7 If M_1 is source multiplicative, μ_2 has no atomic part, $S_1 = \{0,1\}$, $\mu^0_1 \perp \mu^1_1$ and μ^0_1 is completely degenerate at 0, then M_1 is singular.

PROOF All of the mass of μ_2 is collapsed onto zero when multiplying by zero, so that μ^0_3 is degenerate at zero with mass 1 there, while μ^1_3 does not have an atomic part at the origin.

Just to be different we could let S_{2k} rather than S_{2k-1} be the field from which S_{2k+1} is generated. We say that the k-th stage is simple operator multiplicative if:

1. S_{2k+1} is a Banach space
2. S_{2k-1} is S_{2k+1}
3. S_{2k} is isomorphic to the field of S_{2k+1} and if $a \varepsilon S_{2k-1}$, $b \varepsilon S_{2k}$ then $b(a)=b \cdot a \varepsilon S_{2k+1}$

This kind of stage models a simple fading: the amplitude of the signal is a random variable. Another kind of "multiplication", in which the space of operators has a group structure, is a delay stage:

1. S_{2k+1} is a space of functions. Each function has the real line as at least one argument, say the first.
2. S_{2k-1} is S_{2k+1}
3. S_{2k} is a group of translation operators: if $a \varepsilon S_{2k-1}$ and $T_h \varepsilon S_{2k}$ then $T_h : a(t_1, \ldots) \rightarrow a(t_1+h, \ldots)$

These two last types can be combined to give a gross model of multi-path transmission effects. If $a \varepsilon S_{2k-1}$ and $T \varepsilon S_{2k}$ then

$$T(\underline{b},\underline{h})(a)=\Sigma b_i a(t+h_i)$$

A natural question to ask is, what kind of topology can S_{2k} have and still induce a measure on S_{2k+1}. The simple operator multiplication, being a subspace of the space of linear operators, presents no problem. What about the delay stage? It, too, is a linear operation, but now the "natural" norm, $||T_h||=|h|$ does not always induce a suitable topology: the functions in S_{2k+1} must be continuous first.

4. APPLICATIONS TO SONAR

Models of sonar systems must include descriptions of the process of sound transmission, either explicitly or implicitly. Often these descriptions begin by assuming not only that the scalar wave equation is applicable, but that a particular solution of the wave equation can be used. (for example, incident plane waves, Bryn [1]) While this approach avoids a great deal of complexity, the assumptions involved may create serious problems (e.g., singular models) whose origins have been obscured by the lack of explicit detail in the initial model building.

One of the concerns of this chapter is to identify and understand all of the assumptions that are made in the sonar models to be used here, so the chapter starts with a discussion of the basic physics of sound transmission (this follows Sokolnikoff [1] quite closely). Derivation of the scalar wave equation requires several major assumptions which are identified for discussion in section 4.6.

Sections 4.2 and 4.3 are primarily concerned with the kernel of the transmission operator defined implicitly by solution of the inhomogeneous wave equation in 2 and 4 variables, respectively. The kernel of the 2 variable operator can be characterized nicely, chiefly because the solution of partial differential equations in two variables can be reduced to the solution of ordinary differential equations along characteristics. Consideration of the 4 variable transmission operator is much more difficult; it has only been possible to make a start here.

Sections 4.4 and 4.5 are devoted to sonar problems; the theory previously developed is applied to the construction and analysis of sonar models in 1 and 3 spatial dimensions. Finally, section 4.6 contains a summary of the work in the area of singularity of sonar models and suggestions on the construction of non-singular models.

Throughout section 4.1, especially, use is made of the Einstein summation convention: whenever a subscript occurs on the right side of an equation that does not appear on the left, a summation over all values of that index is implied. The range of the indices is from 1 to the dimension of the space in which one is working, normally 2 to 4.

4.1 The Physics of Sound Transmission

Underwater sound transmission is treated here as a special case of wave propagation in a continuous medium. As mentioned above, we want to derive the differential equation which governs sound transmission in order to point out all of the physical assumptions that are implicit in the use of the scalar wave equation to model sound propagation. Since we want to uncover assumptions, we are forced to begin at a general level:

Strains

Consider two states of the same material body:

Initial State		Deformed State
Spatial Region	T_o	T
Reference Frame	Y^o	X
Coordinates of a point P	${}^{i}y$	x^i

As our first assumption, (a physically reasonable one though), let the deformation of T_0 into T be of class C^1 and 1:1 so that the point transformation

$$x^i = x^i({}^1y, {}^2y, {}^3y, t)$$

has an inverse

$${}^iy = {}^iy(x^1, x^2, x^3, t)$$

for all values of the deformation parameter t, and the derivatives

$$\partial x^i / \partial^j y \qquad\qquad \partial^i y / \partial x^j$$

exist and are continuous.

If ds_o^2 and ds^2 are the initial and deformed lengths of an infinitesimal arc, their difference represents the strain produced in the medium by the deformation. Restricting ourselves to rectangular Cartesian coordinates alone, we can write:

$$ds_o^2 = d^iy\, d^iy$$
$$ds^2 = dx^i dx^i$$

so that

$$ds^2 - ds_o^2 = (\partial x^k/\partial^i y \cdot \partial x^k/\partial^j y - {}_{ij}\delta)\, d^iy\, d^jy$$
$$= (\delta_{ij} - \partial^k y/\partial x^i \cdot \partial^k y/\partial x^j)\, dx^i dx^j$$

or,

$$ds^2 - ds_o^2 = 2\, {}_{ij}\eta(y,t)\, d^iy\, d^jy$$
$$= 2\epsilon_{ij}(x,t)\, dx^i dx^j$$

where ${}_{ij}\eta$ is the Langrangian strain tensor and ϵ_{ij} is the Eulerian strain tensor.

Letting $\xi = (\xi^1, \xi^2\ \xi^3)$ be the displacement vector, we can write:

$$\xi^i(y,t) = x^i(y,t) - {}^iy$$
$$\xi^i(x,t) = x^i - {}^iy(x,t)$$

Differentiating and substituting for $\partial x^i/\partial^j y$ and $\partial^i y/\partial x^j$, we see that

$$2{}_{ij}\eta(y,t)=\partial\xi^i/\partial^j y + \partial\xi^j/\partial^i y + \partial\xi^k/\partial^i y \cdot \partial\xi^k/\partial^j y$$

$$2\varepsilon_{ij}(x,t)=\partial\xi^i/\partial x^j \quad \partial\xi^j/\partial x^i - \partial\xi^k/\partial x^i \cdot \partial\xi^k/\partial x^j$$

By assuming infinitesimal strains, we can drop the product terms and also disregard the differences between the initial and deformed coordinates since

$$\partial x^i/\partial^j y=\delta^i_j + \partial\xi^i/\partial^j y = \delta^i_j$$

Having done this, we find

$$\begin{aligned} {}_{ij}\eta=\varepsilon_{ij}&=1/2 \cdot (\partial\xi^i/\partial x^j + \partial\xi^j/\partial x^i) \\ &=(\xi_{i,j} + \xi_{j,i})/2 \qquad \text{(E4.1)} \end{aligned}$$

In this linear theory, ε_{ij} is symmetric.

This set of assumptions is unreasonable in general, but in passive sonar problems the particle displacements are so very small as to justify their adoption.

Stress

Stress, the force per unit of area, is also characterized by a symmetric tensor, the stress tensor τ^{ij}. There are no assumptions hidden under the tensor cover.

Equation of Equilibrium

Consider a body T which is in equilibrium under surface forces given by the stress tensor τ^{ij} and volume forces given by the force per unit of volume F^i. Letting λ_i be a fixed unit vector and n_j be the unit surface normal, the assumed equilibrium is expressed by:

$$\int_V F^i\lambda_i d\tau + \int_S \tau^{ij}\lambda_i n_j d\sigma = 0$$

for every subregion V of T bounded by the surface S. Applying the divergence theorem and noting that λ_i is a constant, we see that, at every point in T

$$F^i + \tau^{ij}_{,j} = 0 \qquad \text{(E4.2)}$$

Equations of Motion

Using D'Alembert's principle, we add the inertial force $-\rho a^i$ to obtain the equations of motion:

$$F^i + \tau^{ij}_{,j} - \rho a^i = 0 \qquad \text{(E4.3)}$$

where ρ is the density and a^i is the acceleration

Stress-Strain Relationships

An additional basic assumption is that stress and strain are linearly related:

$$\tau^{ij} = c^{ij}_{km}\varepsilon^{km}$$

Of the 81 components of the tensor c^{ij}_{km}, 27 are eliminated by the symmetry of τ^{ij}, and 18 more by the symmetry (in the linear theory) of ε^{km}. Since the strain energy per unit volume is

$$\delta W = \tau^{ij}\delta\varepsilon_{ij} = c^{ij}_{km}\varepsilon^{km}\delta\varepsilon_{ij}$$

$$W = c^{ij}_{km}\varepsilon^{km}\varepsilon_{ij}/2 = c^{km}_{ij}\varepsilon_{km}\varepsilon^{ij}/2$$

we see that $c^{ij}_{km} = c^{km}_{ij}$, which eliminates 15 more components.

Isotropic Medium

If the medium is isotropic then more components can be eliminated:

a) 3 interchanges of axes: 18 components eliminated.

b) Together with reversal of the sense of an axis, a) shows that all but three independent coefficients are zero.

c) Invariance under rotation shows, finally, that

$$c_{12}^{12} = c_{11}^{11} - c_{22}^{11}$$

The remaining components can be taken in the form:

$$c_{12}^{12}=c_{13}^{13}=c_{23}^{23}=c_{21}^{21}=c_{31}^{31}=c_{32}^{32}= \quad 2\mu$$
$$c_{22}^{11}=c_{33}^{11}=c_{33}^{22}=c_{11}^{22}=c_{11}^{33}=c_{22}^{33}= \lambda$$
$$c_{11}^{11}=c_{22}^{22}=c_{33}^{33}= \quad \lambda+2\mu$$

from which we obtain the isotropic relationship:

$$\tau_j^i = \lambda\varepsilon_k^k\delta_j^i + 2\mu\varepsilon_j^i \tag{E4.4}$$

The assumption of isotropy is a good one in sonar work.

Isotropic Equations of Motion

Putting E4.4 into E4.3, and using E4.1, we obtain, using D'Alembert's principle again

$$(\lambda+\mu)\xi^k_{,ki} + \mu g^{jk}\xi_{i,jk} = \rho\ddot{\xi}_i \tag{E4.5}$$

No assumptions about λ, μ or ρ have been made. All three could be functions of position. The metric tensor, g^{jk}, would become the Kronecker delta δ^{jk}, in Cartesian coordinates of course.

Perfect Fluid

A perfect fluid is one in which $\mu=0$. If we define the dilation, I, to be the particle divergence,

$$I= \xi^k_{,k} = \nabla\cdot\xi$$

then we have

$$\lambda\xi^k_{,ki} = \rho\ddot{\xi}_i$$

as the equation of motion for an isotropic perfect fluid.

To work the dilation into this equation, differentiate with respect to x^i and sum on i to obtain

$$\lambda \nabla^2 I = \rho \ddot{I}$$

where we have assumed that variations in λ and ρ are small compared with those of I. Unless λ and ρ are taken to be constants, a theory based on this description of sound propagation would not be valid for arbitrarily low frequency waves. The assumption that water is a perfect fluid is more innocuous in sonar work; the effect of this assumption is to make it impossible for the model of the medium to support shear waves. These do exist in water, but the low viscosity limits their range so greatly that they can be safely ignored.

Since the stress tensor for a perfect fluid is

$$\tau^i_j = \lambda I \delta^i_j$$

we can define $p = \tau^i_i$ and $k = -\lambda$ in order to obtain

$$p = -kI$$

and if $c^2 = -k/\rho$ then

$$c^2 \nabla^2 p = \ddot{p} \qquad \text{(E4.6)}$$

That is, within the limitations of our assumptions, (the important ones being: very small particle displacements; a linear stress-strain relationship; nearly constant bulk modulus and density, yielding a nearly constant speed of sound) sound propagation obeys the scalar wave equation.

It would be very interesting to develop the following sonar models with these assumptions weakened or eliminated, but it does not seem possible to do this.

4.2 The Two Variable Wave Equation

The scalar wave equation,

$$\nabla^2 p = \ddot{p}/c \qquad \text{(E4.6)}$$

describes the propagation of sound when wave amplitudes are small and c, the speed of cound, is nearly constant. While this classic differential equation has received the attention of many minds for well over 100 years, its properties are still not completely known and cataloged. (Courant [1])

In this section, some interesting properties of the equation in two variables, time and one spatial dimension, are developed, using very elementary methods, based on the properties of characteristics for 2 variable differential equations. In section 4.3, following, similar results are sought for 4 variables, but here a variety of methods must be employed, none with complete success. The results sought are essentially concerned with the size of the null space of the operator defined by (E4.6) when operating on Cauchy initial data, with and without inhomogeneous terms.

Consider the differential operator#

$$W[p] = c^2(x,t)p_{xx} - p_{tt} \qquad \text{(E4.7)}$$

For scalar wave phenomena, c^2 is positive so that W is everywhere hyperbolic. The characteristic curves for W, those along which W

\# The prerequisites for this discussion can be found in any standard text, e.g. Courant [1] or Morse [1]. Note that independent variables, x,y,z,t, used as subscripts denote differentiation: $\partial f(\cdot)/\partial x = f_x(\cdot)$. Other subscripts, such as n,k etc., do not denote differentiation, but serve merely as auxiliary arguments to functions.

is an interior operator, are defined by the equation

$$c^2\phi_x{}^2 - \phi_t{}^2 = 0$$

where $\phi(x,t)=0$ defines a curve in the (x,t) plane which we assume to be regular, i.e., ϕ_x and ϕ_t may not both vanish simultaneously.

We see that $c\phi_x = \pm\phi_t$ on the characteristics, or, for constant c, $\phi = x \pm ct + x_o$. Characteristics are also rays, or, the directions of propagation of wavefronts (Figure 4.1).

Now consider a simple radiation problem for constant c:

$$W[p] = -c\cdot S_t(t)\delta(x=a)$$

$$p(x,0) = 0 \qquad x\neq a$$

$$p_t(x,0) = 0 \qquad x\neq a$$

where $S(t)$ is the intensity of a point source located at $x=a$. The solution, $p(x,t)$ for all $t>0$ is given by

$$p(x,t) = S(t-|x-a|/c) \tag{E4.8}$$

That is, the signal produced by the point source divides in half, one half propagates to the left, $x<a$, the other to the right, $x>a$. The propagation is along the rays. If many point sources are present, or distributed sources are assumed, then E4.8 generalizes to

$$p(x,t) = (1/2)\int dS_\lambda(t-|x-\lambda|/c) \tag{E4.9}$$

where dS_λ is the signal intensity at $x=\lambda$, and the integration is in the Riemann-Stieltjes sense. If S_λ vanishes for all $\lambda<=\underline{a}$, then in the region $x<\underline{a}$,

$$p(x,t) = S_{\underline{a}}(t-|x-\underline{a}|/c)/2 \tag{E4.10}$$

where $S_{\underline{a}}$ is an equivalent source defined by

$$S_{\underline{a}}(t) = \int dS_\lambda(t-|\underline{a}-\lambda|/c)$$

An analogous introduction of an equivalent point source is possible if S_λ vanishes for $\lambda>\overline{a}$.

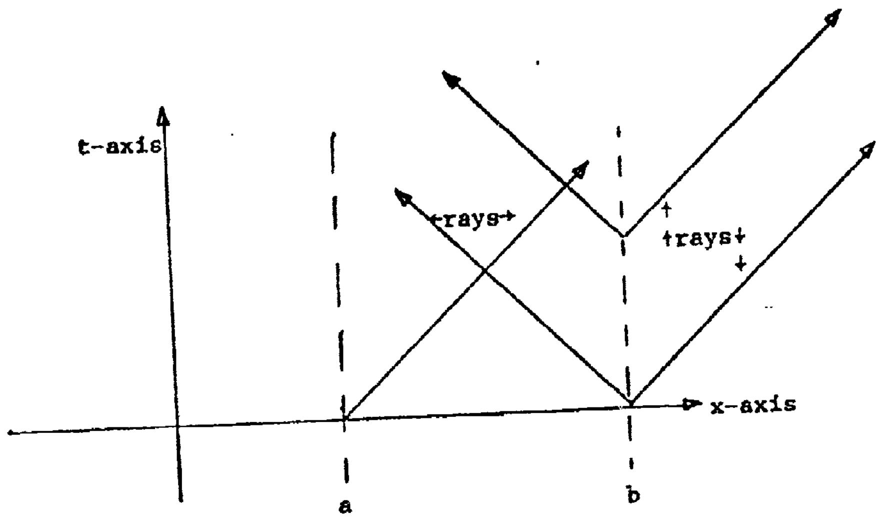


Figure 4.1

Constant Wave Velocity Rays

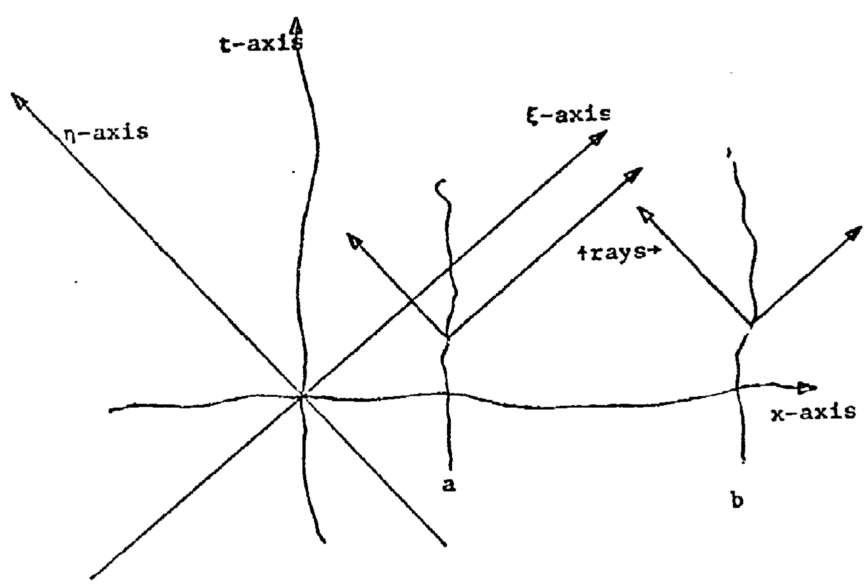


Figure 4.2

Variable Wave Velocity, Transformed Variables

Now we ask the question: when are different source distributions indistinguishable to an observer? That is, when do two different source distributions differ only by a vector which lies in the nullspace of the linear operator $T:S\to p$ defined implicitly by $W[p]= -cS_t$? For the observer we will take an open ray-connected set D of the (x,t) plane, and by ray-connected, we will mean that any two points $r,s\epsilon D$ can be connected by a ("zig-zag") path of ray-segments. Let K_D be the set of constant functions on D. Letting D_a be the set of waveforms from a point source a that are observable in D, we see that

THEOREM 1 If D is an observation region, S_a and S_b are point sources, and $a<D<b$, then $D_a \cap D_b =0$.

PROOF Let $w(x,t)\epsilon D_a \cap D_b$ and take any two points (x_o,t_o) and (x_1,t_1), connected by the ray-path Γ. $w(x,t)$ is a constant along each ray segment of (positive,negative) slope since $(w\epsilon D_a,\ w\epsilon D_b)$, hence w is constant along all of Γ so that $w(x_o,t_o)=w(x_1,t_1)$. Since the points were arbitrary, $w\epsilon K_D$. Since neither D_a nor D_b contain K_D, the theorem is proved. (constant functions are not in D_a or D_b because they violate the initial conditions)

This result is not unexpected, and admits obvious generalizations to distributed sources and sources within the observation region. A generalization to cover random sound velocities is also possible, although not so obvious (Figure 4.2).

THEOREM 2 If D is an observation region, S_a and S_b are point sources with $a<D<b$ and transmission is governed by

$$W[u] = c^2u_{xx}-u_{tt}$$

with zero initial conditions, and

$$0<\underline{c}\le c(x,t)\le\bar{c}<\infty$$

and $c(\cdot,\cdot)$ twice continuously differentiable, then

$D_a \cap D_b = 0$.

PROOF Make a change of coordinates in the equation

$$W[u] = c^2(x,t)u_{xx}-u_{tt}=h(x,t)$$

by letting

$$\xi=\xi(x,t) \qquad \eta=\eta(x,t) \tag{E4.11}$$

where we assume that the transformation $\psi:(x,t)\to(\xi,\eta)$ is everywhere invertible, i.e., $\xi_x\eta_t-\xi_t\eta_x\neq 0$. We find:

$$u_{\xi\xi}[c^2\xi_x^2-\xi_t{}^2] + u_{\xi\eta}[c^2\xi_x\eta_x-\xi_t\eta_t] + u_{\eta\eta}[c^2\eta_x^2-\eta_t{}^2]$$
$$=h(x(\xi,\eta),t(\xi,\eta)) = g(\xi,\eta) \tag{E4.12}$$

If the transformation $\psi:(x,t)\to(\xi,\eta)$ is chosen so that

$$c\xi_x=\xi_t \qquad c\eta_x=-\eta_t \tag{E4.13}$$

(solution of these two first order partial differential equations is certain since c is in C^2) then the transformed equation, E4.12, becomes

$$2\xi_t\eta_t u_{\xi\eta}=g(\xi,\eta)$$

Since c is bounded away from zero, ξ_t can be zero only if ξ_x is zero (E4.13). But if either is, both are, and the transformation becomes singular, which has been ruled out. Put another way, solutions of E4.13 which are bounded away from zero exist since c is bounded away from 0; see Courant [1], page 491 et. seq.

As a result, E4.12 can be written as

$$u_{\xi\eta} = g(\xi,\eta)/2\xi_t\eta_t \qquad \text{(E4.14)}$$

This is a canonical form for the constant coefficient wave equation; ξ=constant and η=constant are characteristics and rays of the solution. Theorem 1 now applies showing that

$$D_a(\xi,\eta) \cap D_b(\xi,\eta) = 0$$

and the theorem is established.

It is now easy to prove that random wave velocities do not enlarge the nullspace of the transmission operator very much:

THEOREM 3 If D is an observation region, S_a and S_b are point sources and transmission is governed by

$$W_c[u] = c^2(x,t)u_{xx} - u_{tt}$$

with zero initial conditions, and c is randomly chosen from a set C* each of whose elements c satisfy

$$0 < \underline{c} <= c(x,t) <= \bar{c} < \infty$$

and each of which is twice continuously differentiable, then $D_a \cap D_b = 0$. (by D_a we mean $\bigcup_{c \in C^*} D_a(c)$

PROOF $D_a(c) \cap D_b(c) = 0$ by Theorem 2.

This result shows that, with only two discrete directions from which signals can come, random propagation can not hurt your discrimination. A random medium just is not able to make a left going wave look like a right going one.

4.3 The Four Variable Wave Equation

Now let W be the differential wave operator in 4 variables:

$$W[p]=\nabla^2 p-p_{tt} \qquad \text{(E4.15)}$$

We are interested in an inverse problem in differential equations: given suitable x and the equation $W[p]=x$ with suitable boundary conditions, we can solve for p. But, given p, can we solve for x? The obvious answer is yes: everywhere p is known, x is known also, just apply W to p. But can x be determined in a region of the independent variables removed from the region in which p is known? In general the answer is no. With further assumptions about x, however, it can be yes.

This whole area of inverse problems in parital differential equations is difficult and relatively untouched: the problems are generally ill-conditioned. If we think in terms of some inverse operator $J:p\to x$ then J will be unbounded. What little that has been done with these problems has been done with relatively tractable equations: Poisson's equation, where the entire apparatus of complex variable theory can be used, and equations in two independent variables where reduction to ordinary differential equations along characteristics is a powerful trick. (Lavrentiev [1])

A complete treatment of this problem can not be expected, then. The most we can hope to do is to treat a few special cases and to acquire some insight.

Looking at Figure 4.3, we see that at each point $V=(t,x,y,z)$ in $\underline{R}^4$, a characteristic conoid exists for the operator W. This conoid is a right circular 45^o hypercone, axis parallel to the t-axis. Let P_o be a hyperplane $t=t_o>t_V$ and P_g a hyperplane $t=t_g<t_V$. Intuitively,

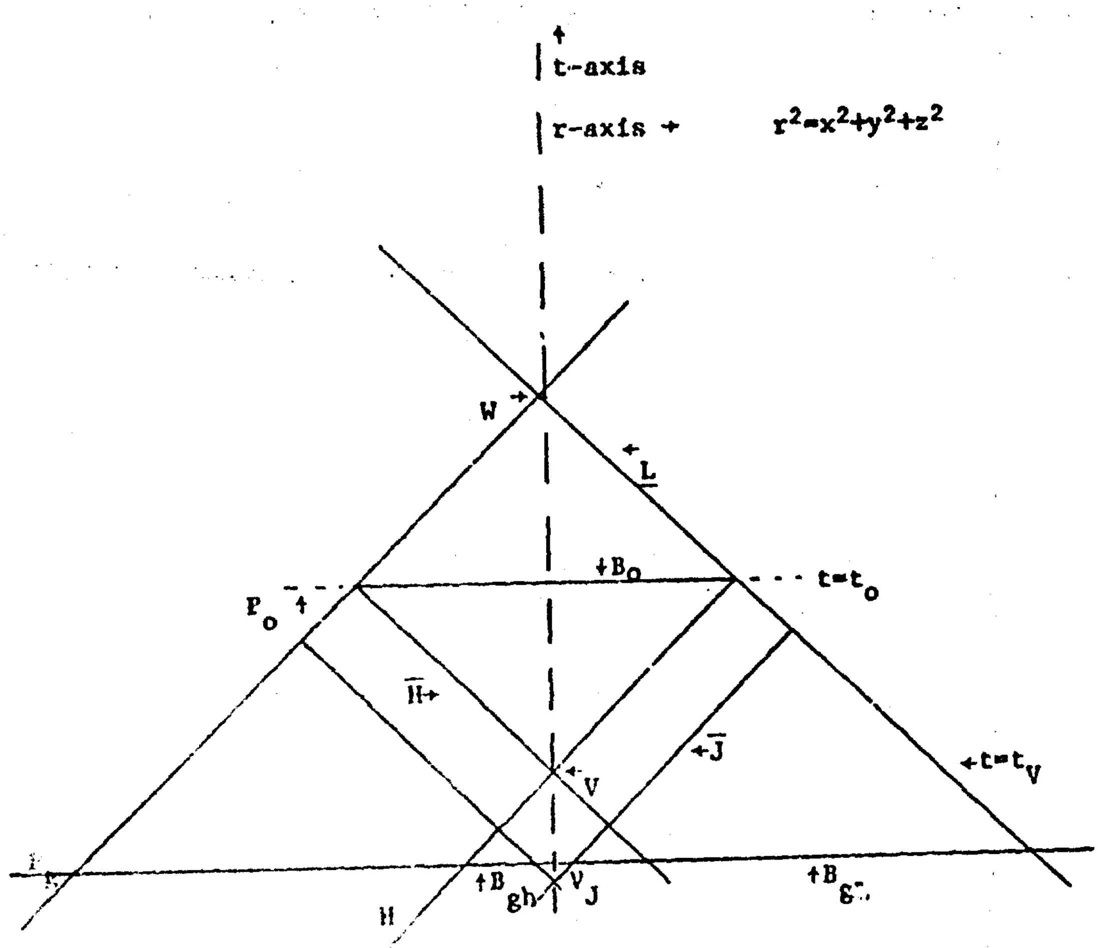


Figure 4.3

Projection of Characteristic Surfaces onto the (r,t) Plane

[illegible]: complete cone, vertex at V

H: lower sheet of cone, $t<t_V$

$\bar{H}$: upper sheet of cone, $t>t_V$

P_o is the observation plane, and P_g is the given plane, or the plane of assumed values. The conical frustrum whose base is the intersection, B_o, of P_o and H is mirrored by reflection across P_o into a frustrum of another right circular 45^o hypercone, L, which has a vertex W in the plane $t=t_L=2t_o-t_V$. The sheet of L, called $\underline{L}$, which lies below t_L intersects the plane P_g also; we call the base thus determined B_{gL}, while the base of H in the plane P_g is B_{gh}.

The uniqueness (but not the existence) of any solution to the initial value problem

$$W[u] = 0 \qquad \text{in } \underline{L} \cap \overline{H}$$

$$(u,u_t) = (\psi_o,\psi_1) \qquad \text{on } B_o$$

is established by the following well known theorem. The proof, based on an "energy" integral, is worth repeating for the light it sheds on the behavior of the wave equation.

THEOREM 4 A solution u to the differential equation $W[u]=0$ in $\underline{L} \cap \overline{H}$ satisfying arbitrary initial conditions

$$(u,u_t)=(\psi_o,\psi_1)$$

on B_o (where ψ_o is twice continuously differentiable and ψ_1 is once continuously differentiable) is unique.

PROOF We will show that the initial conditions uniquely determine the value of the solution at the point W. As every other point within $\underline{L}\cap\overline{H}$ is the vertex of a characteristic cone whose base is within the sphere B_o, this will establish the uniqueness of the solution throughout $\underline{L}\cap\overline{H}$. Suppose, then, that u_1 and u_2 are both twice

continuously differentiable solutions of $W[u]=0$, with the same initial conditions $u_1=u_2=\psi_o$, $\dot{u}_1=\dot{u}_2=\psi_1$ on B_o. Their difference, $u=u_1-u_2$, must be everywhere zero if uniqueness is to obtain. Noting that u satisfies $W[u]=0$ within $\underline{L}\cap\overline{H}$ with zero initial data on B_o, we integrate over all of $\underline{L}$ above B_o, that is, K:

$$0=\int_K \partial u/\partial t(\partial^2u/\partial t^2 - \nabla^2u)d\underline{x}dt$$

since

$$\partial u/\partial t \cdot \partial^2u/\partial t^2 = (\partial/\partial t)(\partial u/\partial t)^2/2$$

and

$$\begin{aligned}\partial u/\partial t \cdot \partial^2u/\partial x_i^2 &= (\partial/\partial x_i)(\partial u/\partial t \cdot \partial u/\partial x_i) - \\ &\qquad \partial u/\partial x_i \cdot \partial^2u/\partial t\partial x_i \\ &= (\partial/\partial x_i)(\partial u/\partial t \cdot \partial u/\partial x_i) - \\ &\qquad (\partial/\partial t)(\partial u/\partial x_i)^2/2\end{aligned}$$

we see that the above integral can be written as

$$\begin{aligned}0=\int_K\{(\partial/\partial t)[(\partial u/\partial t)^2+ \sum_i(\partial u/\partial x_i)^2]/2 - \\ \sum_i(\partial/\partial x_i)(\partial u/\partial t \cdot \partial u/\partial x_i)\}d\underline{x}dt\end{aligned}$$

Since Gauss' theorem says $\int_V \nabla\cdot F dv = \int_S F\cdot\underline{n}d\underline{s}$, we have:

$$\begin{aligned}0=(1/2)[\int_{B_o}d\sigma + \int_{K'} d\sigma]\{[(\partial u/\partial t)^2+ \sum_i(\partial u/\partial x_i)^2]\cos(n,t) \\ -2\sum_i(\partial u/\partial t)(\partial u/\partial x_i)\cos(n,x_i)\}\end{aligned} \tag{E4.16}$$

where K' is the lateral surface of K. The integral over B_o is zero since the initial data are zero. Multiplying and dividing by $\cos(n,t)$, which is constant on K', and using the identity, good on K',

$$\cos^2(n,t) = \sum\cos^2(n,x_i) \tag{E4.17}$$

we obtain

$$0=(1/2\cos(n,t))\int_{K'} \Sigma_i[(\partial u/\partial t)\cos(n,x_i)-(\partial u/\partial x_i)\cos(n,t)]^2 d\sigma$$

From this it follows that on K'

$$u_t/\cos(n,t) = u_{x_i}/\cos(n,x_i) = v$$

Using this fact, we evaluate the change in u along some generator, m, of K:

$$\begin{aligned}\partial u/\partial m &= u_t\cos(m,t) + \Sigma u_{x_i}\cos(m,x_i)\\ &=v[\cos(n,t)\cos(m,t)+ \Sigma\cos(n,x_i)\cos(m,x_i)]\\ &=v\cos(m,n)\end{aligned}$$

which is zero since a generator and normal are perpendicular. Now, letting the generator m meet B_o in the point m_o,

$$\int_{m_o}^{W} (\partial u/\partial m)dm = 0 = u(W)-u(m_o)$$

and since $u(m_o)=0$, so does u(W), which establishes the theorem.

This theorem can be interpreted in different ways. From the usual point of view it proves the uniqueness of a linear operator W which maps $(\psi,\psi_t)\varepsilon C^2(B_o)xC^1(B_o)$ into $u\varepsilon C^2(\underline{L}\cap\bar{H})$. Applying the theorem to the sphere B_g, a more symmetric view has us consider the linear operator V which maps $(\psi,\psi_t)\varepsilon C^2(B_g)xC^1(B_g)$ into $(u,u_t)\varepsilon C^2(B_o)xC^1(B_o)$. Since V maps functions on a given domain into functions with a smaller domain, intuition says that V should be many to one, i.e., have a non-vanishing kernel. If this were not the case, then the following theorem would hold:

Conjecture 1   A solution to the differential equation $W[u]=0$ in L satisfying the conditions

$$(u,u_t) = (\psi,\psi_t) \quad \text{on } B_o$$

with $\psi \in C^2$, $\psi_t \in C^1$ and $u=0$ in $\underline{H}$, is unique.

Attempted Proof: We will attempt to prove uniqueness by considering the difference, u, of the two solutions u_1 and u_2 with the same data on B_o. Since $\underline{L}$ extends to $-\infty$, we will introduce a hypercone J with upper sheet $\overline{J}$ and axis r=0, vertex V_J at some point below V, (Figure 4.3) and consider u only within the region $\underline{L}\cap\overline{J}$. Within $\underline{L}\cap\overline{H}$ uniqueness follows directly from Theorem 4. Within $J\cap H$, u is zero by assumption. This leaves a region $L\cap J - H$, bounded on the outside by portions $\underline{L}'$, $\overline{J}'$ of the sheets of $\underline{L}$ and $\overline{J}$, on the inside by portions $\underline{H}'$ and H'' of the upper and lower sheets of H. The steps in the proof of Theorem 4 carry over for this region, up to [illegible], which appears here with 4 surfaces instead of 2:

$$0 = [illegible] [\int_{L'} d\sigma + \int_{\overline{J}} d\sigma + \int_{H'} d\sigma + \int_{\underline{H}'} d\sigma](\ldots)$$

The integrals over $\overline{H}'$ and $\underline{H}'$ vanish since u is zero [illegible] before, we multiply and divide by cos(n,t), and apply the identity E4.17 to find:

$$0 = [illegible] [\int_{L'} d\sigma + \int_{J'} d\sigma](\ldots)$$

The two [illegible] have opposite signs because one (+ sign) [illegible] outgoing waves across $\underline{L}'$, while the other [illegible] waves across $\overline{J}'$. Thus fails the proof.

This failure of an attempted proof does not show that the conjecture is false, but it strongly suggests that it is. Note that this proof can be carried through successfully in one spatial dimension because ingoing/outgoing translates into leftgoing/right-going in one spatial dimension, and these wave types are independent. The proof also is successful in 3 dimensions if u is restricted to be spherically symmetrical, for in that case, ingoing/outgoing waves types are independent.

The initial conditions on B_g that lie in the kernel of the operator V are still unknown, but, it seems that there are some. Furthermore, this kernel of V is in addition to the kernel consisting of functions which are zero on B_g-B_{gh}. These are in the kernel of V because Huyghen's principle holds in 3 spatial dimensions, so that B_g-B_{gh} is the complete domain of dependence of B_o or the point W. In two spatial dimensions (in fact, all even spatial dimensions) by contrast, the failure of Huyghen's principle makes V even less invertible.

While this discussion has been about V, an operator from initial conditions into initial conditions, by Duhamel's principle the same will be true of the transmission operator that maps into u from source distributions. That is, the transmission operator will have a kernel whenever u can only be observed in a limited region. In order to display one way in which this can happen, we look at the field produced by a spherical shell source $h(r,t)=h(t)\delta(r-s)$. Green's function in the transform domain is

$$g_\omega(r,r_o)=\exp[i\omega R]/4\pi R$$

$$R = |r-r_o|$$

so the field at r_o is

$$U(r_o,s,\omega) = \int dv H(\omega)\delta(r-s)\exp[i\omega R]/4\pi R$$

where $H(\omega)$ is the Fourier transform of $h(t)$. A little algebra gives:

$$U(r_o,s,\omega) = sH(\omega)\exp[i\omega s]\sin[\omega r_o]/\omega r_o$$

From this we see that exact cancellation is possible everywhere ($r_o<s$) inside the spherical shell source if only a second spherical shell outside the first, at a radius $q>s$, is excited by

$$G(\omega) = H(\omega)\exp[i\omega(s-q)]s/q$$

This cancellation depends upon the spherical symmetry. To see just how critical this dependence is, consider two hemispherical sources at radii s and q. One finds that the field from one hemisphere is:

$$U(r_o,s,\omega) = sH(\omega)\{\exp[i\omega(r_o^2+s^2)^{1/2}]-\exp[i\omega(r_o-s)]\}/2i\omega r_o$$

and for both it is:

$$2i\omega r_o[U_H-U_G] = sH(\omega)\exp[i\omega(r_o^2+s^2)^{1/2}] - sH(\omega)\exp[i\omega(r_o-s)]$$
$$-qG(\omega)\exp[i\omega(r_o^2+q^2)^{1/2}] + qG(\omega)\exp[i\omega(r_o-q)]$$

Proper choice of $G(\omega)$ can cause cancellation of the second and fourth terms, but, no cancellation of the first and third terms is possible. Some leakage around the edges of the hemispheres always occurs.

If the wave equation is written for the velocity potential ψ, where particle velocity is $\underline{u}=-\nabla\psi$ and pressure is $p=\rho\psi_t$ for a density ρ, then the energy flux, or intensity is

$$I=-\rho\psi_t\nabla\psi=p\underline{u} \tag{E4.18}$$

Given ψ in some region, what cancels it? Obviously, it is $-\psi$. But,

$$I(\psi) = I(-\psi)$$

That is, the wave energy flux is the same for ψ and $-\psi$. In words,

the cancelling wave is going in exactly the same direction at every point of cancellation. This explains the leakage discovered in the two hemisphere problem; failure of spherical symmetry at the edges of the hemispheres results in the generation of waves that are not going in exactly the same direction.

Now let a point source launch a wave, ψ, and follow a portion of the wave front that travels towards the origin. This portion of the wavefront can be cancelled, but only be another wavefront traveling the same direction but with opposite sign: $-\psi$. Such a wavelet could only be generated by a (portion of) a spherical shell with the point source as center. In particular, a spherical shell of large radius with center at the origin can not launch such a wave.

Summarizing these bits and pieces of evidence, we state

Conjecture 2 Let u be determined by

$W[u]=h(t)\delta(\underline{\gamma}=\underline{\gamma}_0) + g(\theta,\phi)\delta(r=r_o)$

$(u,u_t) = (0,0)$ at $t=-\infty$

$h(t)$ a given point source at the point $\underline{\gamma}_0\neq 0$

$g(\theta,\phi)$ a given spherical shell source at radius r_o.

Then u is non-zero in any open connected set containing the origin.

This conjecture requires proof from a mathematical point of view. From an engineering viewpoint, it can be regarded as true. Extensions to a finite sumber of point sources and non-concentric shells follow from the truth of this conjecture.

4.4 Passive Sonar in One-space

Following the lines of section 2.4 we begin with the construction of a class, call it OSD, of level 4 C/D models. Models in this class differ only in their measures, and the class is broad enough to encompass a wide variety of specific sonar models.

Space	Meaning and Description
S_1	{0,1}, standing for {no signal,signal} as usual.
S_2	Encoding operators e_h with values in S_3 defined by $e_h(a) = ah$ The set of encoding operators is isomorphic and isometric to S_3 under the map $\psi: e_h \to h$.
S_3	Source distribution functions for the plane, representing the signal and forming a subspace of S_5. These are limited by the assumption that $h(x,t)=0$ for all $x>\alpha$, that is, the signal is confined to a right half-plane.
S_4	Additive noise operators, $n_g(h)=g+h$ with values in S_5. The noise operators are isomorphic and isometric to a subspace of S_5 under the map $\psi: n_g \to g$. This subspace is restricted by the assumption that $g(x,t)=0$ for all $x<\beta$, so that the noise is restricted to a left half plane.
S_5	Source distribution functions for the plane. These functions are intended to lie within the domain of the transmission operators in the space S_6, but since they enter via an integral, no great number of restrictions need be placed upon them. We choose the space of all Lebesgue square integrable functions in all space-time.

Space	Meaning and Description
	Intuitively, this space, $L^2(\underline{R}^2)$, is reasonable, representing as it does a finite energy constraint on the signals and noise.
S_6	Transmission mappings, T_c, from S_5 into S_7, defined implicitly by solution of $c^2(x,t)u_{xx}-u_{tt}=h(x,t)$ where $c(x,t)$ is twice continuously differentiable and $0<\underline{c}<=c(x,t)<=\overline{c}<\infty$ and $h\varepsilon S_5$. The Cauchy initial data are prescribed as $(u,u_t)=(0,0)$ on the line $t=-\infty$.
S_7	All solutions of the inhomogeneous wave equation, but considered in a given observation region D with $\alpha<D<\beta$ S_7 is a linear manifold within $L^2(D)$ in view of S_5.
S_8	Detection operators. All possible measurable maps of S_7 into S_9. Which ones are chosen depends upon the particular model of class OSD which is being considered.
S_9	{0,1} standing for {noise alone, signal plus noise}.

We now examine the class OSD for singular models. If M is one of the factorable models in OSD, then stage M_1 is singular by theorem 3.7, while stage M_2 is singular (i.e., preserves singularity) by theorem 3.1. Stage M_3 is also seen to be singular by application of theorem 3.5 combined with theorem 4.3. This leaves only the fourth stage, the detector, between us and singularity of the entire model.

But any detector that maps the support of μ_7^1 into $i\epsilon S_9$ preserves singularity, and there are many of these. Our conclusion, then, is that all of the factorable models in OSD that have "good" detectors, which certainly includes any optimal (Baysian) detectors, are singular, hence inadequate as analytic guides to reality. This holds for quite general signal and noise locations (as long as they are disjoint and the detector is between them) and for arbitrary sound velocities, even random velocities (as long as they are bounded away from zero, and glossing over the inadequacy of the scalar wave equation's description of sound transmission when the sound velocity is not slowly varying, which fact really means that some of the models in the class OSD are not good images of reality, singular or not). Such generality is possible because of the geometric simplicity of the problem: detection really reduces to a determination of the direction of the incident waves: left-going means [illegible] going means noise.

4.5 Passive Sonar in Three-space

Relatively little in the definition of the class OSD needs to be changed in order to generate a class, TSD of 4 stage C/D models applicable to sonar problems in 3 spatial dimensions. We have:

Space	Meaning and Definition
S_1	{0,1}, standing for {no signal, signal}
S_2	Encoding operators e_h with values in S_3 defined by $e_h(a) = ah$. The usual identification with S_3 is provided by the map $\psi: e_h \to h$.
S_3	Source distribution functions for $\underline{R}^4$, representing the signal and forming a subspace of S_5.
S_4	Noise operators, $n_g(h)=g+h$. The usual embedding of S_4 into S_5 is provided by $\psi: n_g \to g$. Notice that S_3 and S_4 have not been restricted to half-spaces as they were in the class OSD.
S_5	Source distribution functions for 4-space, equal to $L^2(\underline{R}^4)$. The discussion of S_5 in the definition of OSD applies here as well.
S_6	A single transmission operator, $T: h \to u$ as defined by solution of the problem $\nabla^2 u - u_{tt} = h(z,t)$, $(u, u_t) = (0,0)$ at $t = -\infty$

The theory for the class TSD is not as comprehensive as that for OSD, as the degenerate nature of S_6 indicates.

Space	Meaning and Definition
S_7	Defined by solution of the wave problem above. A linear manifold within $L^2(D)$ when restricted to a compact set $D\epsilon \underline{R}^4$.
S_8	Detection operators, all possible measurable maps of S_7 into S_9, but with an implicit statement of the observation domain. This represents a slight change from S_8 of class OSD. There, the observation domain was explicit in the definition of S_7, rather than implicit in the choice of a detection operator from S_8.
S_9	{0,1} for {noise alone, signal plus noise}

Each model in TSD is singular in stage M_1 by theorem 3.7, and each factorable model for which suppμ_3 and suppμ_4 are disjoint is singular in stage M_2 by theorem 3.1. Since S_7 contains waves defined throughout $\underline{R}^4$, stage M_3 is singular whenever M_2 is by application of theorem 3.5. This leaves the detector between us and singularity of the model. (notice that this chain of reasoning differs from that used in discussing the singularity of OSD. There stage 3 was key, and singularity occurred when the detection region was between the signal and noise sources. Here we are pushing the key problems back to stage 4, the detection stage.)

When μ_4 peaks up to one on spherical shell sources (at fixed or variable, known or unknown, radii) and μ_3 peaks up to one on point sources (fixed or variable location, one or any finite number of sources) located inside the shell sources, but not at the centers of the shell sources, and the detector selected from S_8

observes an observation region containing the origin, then, if the detector is optimal the model is singular according to conjecture 4.2.

Some of the models considered by Vanderkulk [1] (those without self-noise) are members of the class TSD. M_1 is singular, as ever. μ_3 peaks up on a single point source at infinity, while μ_4 peaks up on a shell source of infinite radius. S_3 and S_4 are effectively separate linear spaces so that M_2 is singular, (theorem 3.1), not by virtue of the spatial separability of the sources, but, because μ_4 generates a process of independent (spatial) increments on the surface of the shell. The μ_3 generated process of point sources has μ_4 measure zero since all of the rest of the sphere has zero excitation. M_3 is singular again, which brings us to the detector. As the number of phones in the Vanderkulk model increases to ∞, observation becomes continuous and the whole model becomes singular, as he shows. This result supports conjecture 4.2.

A fruitful way of looking at these results is this: an optimal detector fed continuous observations can form a zero-width beam pattern, and perform perfect range discrimination for point sources (any finite number of them). The key is the exactly known wave-front available from the source(s). Other signal models that provide exactly known wave-fronts will likewise be singular in the limit of continuous observation.

The introduction of self-noise at the hydrophones is not a cure for this singularity (Vanderkulk [1]). A slight modification of the class TSD suffices to include the self-noise. Spaces S_8 and S_9 of TSD become S_{10} and S_{11} of TSD', and new spaces are introduced:

Space	Meaning and Definition
.	
.	
.	
S_7	Solutions of the wave equation. Subspace of S_9 after restriction to the observation set D.
S_8	Self-noise injection, $d_k(h)=k+h$. S_8 is embedded into a linear subspace of S_9 by the map $\psi: d_k \to k$.
S_9	$L^2(D)$ where D is a compact observation set in $\underline{R}^4$
S_{10}	(old S_8)
S_{11}	(old S_9)

The self-noise, stage 4 of TSD', fails to remove the singularity because it is spatially white, its power spread equally over all of S_9, so, it has zero power on any one dimension of S_9. (in the limit of continuous observation).

4.6 Implications

The general conclusion to be drawn from all of this is that the available sonar models are inadequate. How can they be improved? The possibilities are: stages 1 and 2 might be modified to cause the signal and the noise to overlap, but, this means putting power from the noise at the same spatial locations as the signal, and since the signal can be anywhere, the whole volume of 3-space must be filled with noise sources.# At the same time, the signal must be made into a distributed source in order to destroy the perfect wave-front generated by a point source. This medicine seems excessively bitter -- the noise model that results has little resemblence to the noise sources that we think are present in the ocean. Furthermore, point source signals should be permissible since any distributed source of finite dimensions looks like a point source as it recedes.

As we have seen, introduction of self-noise does nothing for us, so, as the only remaining possibility, the transmission stage must be modified. The cure is easy to talk about, difficult to use. It consists of modeling the randomness of the medium. If there is a low frequency signal cut-off, this modeling can be

Even this might not work since, in the Gaussian noise case with independent radiators, a finite radius spherical ensemble produces the same correlation function within the sphere as does a spherical surface ensemble (Cron [1] [2]). There is also the problem of infinite energy: if an infinite radius spherical volume ensemble is postulated, there must be zero energy generated in every differential element of volume!

accomplished by perturbing the speed of sound in the wave equation since volume inhomogeneities can then be assumed to be much greater than a wavelength (section 4.1). T_λ, the transmission operator perturbed by a small amount λ, maps x into u as defined by solution of

$$(c + \lambda(z,t))^2\nabla^2 u - u_{tt} = x$$

$$(u,u_t) = (0,0) \quad \text{at } t = -\infty$$

Unfortunately, it is difficult to estimate the realism of the low frequency cut-off assumption. Officer [1], for instance, gives estimates for when the eikonal equation is a good approximation to the wave equation, but not for when the wave equation is a good approximation to the physical situation. It is probably not too good, especially in the upper ocean where a fair amount of sea-life serves to complicate things by creating smaller scale volume inhomogeneities.

In order of decreasing realism, and increasing analytic ease, it is suggested that models be modified to

1. Contain the transmission operator defined by the wave equation with random speed of sound and (scattering type) terms due to low frequency signals.
2. Contain the transmission operator defined by the wave equation with random speed of sound, without scattering terms.
3. Contain wavefront perturbation noise introduced as arrival time jitter at each hydrophone.

Only 3 seems simple enough to lead to analytic results. However, extensive analytic and numeric work with 1 and 2 should be done to model the statistics of the arrival time jitter process. This kind of modification of sonar models should have a significant effect on the results of sonar analyses. It is probable, for instance, that a point source of interference will cost considerably more than one hydrophone to null (Schultheiss [2]) when perfect wave-fronts are eliminated. It is not clear what effect a model with jitter will have on detection in the limit of continuous observation. Modeling jitter in that situation should provide an interesting mathematical challenge, as it would seem to require a stochastic process whose elements are (continuous?) maps of a set into itself.

5. SUMMARY

After providing a summary of the contributions made by each of the previous chapters, we present a short list of further research topics (of all sizes).

5.1 Contributions of this Work

The chief contributions made in this work may be divided conveniently along chapter lines:

Chapter	Contribution
1	Analysis of array design problems, showing their relationship to model singularity.
2	Development of a means for classifying most models of communication and detection. Presentation of an adequate and precise definition of model singularity.
3	Discovery of an underlying feature of singularities in certain kinds of models, namely, inequality of the signal and noise subspaces.
4	Application of these results to sonar detection models, with the conclusion that models currently used are inadequate and may give misleading results. Suggestions for improving the models are given.

5.2 Possible Directions for Further Work

The possibilities presented for further work are also conveniently treated on a chapter, section basis: (we restrict this list to problems directly suggested by the work presented here)

Section	Suggested Extensions or Modifications
1.3	Any extension of the theorems on orthogonality versus equivalence of Gaussian measures, as revealed by properties of their covariances, to the sonar case would be very interesting.
2.2	The model apparatus defined in this section provides a convenient skeleton for a taxonomy of detection models, the compilation of which would serve to consolidate the understanding of C/D problems that has been achieved so far and prepare a base to support further achievement.
2.3	Sufficient conditions for existence of induced measures are needed.
2.4,5,6	Further examples could profitably be investigated. N dimensional Gaussian processes, as well as processes defined more directly by their sample spaces await treatment (Parthasarathy [1]). Additional topologies might be investigated.
2.7	Additional performance criteria could well be investigated for continuity properties: Neyman-Pierson, for instance, is closely related to the Baysian risk. Maximum information transfer is another candidate for

Section	Suggested Extensions or Modifications
	investigation. Also, while continuity of Π_C in P and P' has been shown, a proof of continuity in P_H is lacking.
4.3,5	A good deal of open ground lies here, but it may continue to lie fallow through infertility. At any rate, conjecture 4.1 could use a counterexample or a proof while conjecture 4.2 needs a proof, and, many theorems similar to conjecture 4.2 need to be investigated (they would differ from conjecture 4.2 chiefly in the source geometries assummed). This can be paraphrasedby saying that a much deeper understanding of the transmission operator is a prerequisite to better understanding of sonar models.
4.6	Jitter models beckon. Also, analytic work, numeric and field experiments to provide the statistics of the jitter process. Analytic work to justify the jitter model and determine its properties in the limit of [illegible] observation.

Best Available Copy

BIBLIOGRAPHY

Anderson, V.C.

1 "Digital Array Phasing", JASA, 32 (1960) 867

Aronszajn, N.

1 "La Theorie des Noyaux Reproduisants et ses Applications. Premiere Partie", Proc Cambridge Phil. Soc. 39 (1943) 133

2 "Theory of Reproducing Kernels", Trans. Amer. Math. Soc. 68 (1950) 337-404

Bellman, R.

1 Dynamic Programming, Princeton University Press, Princeton, 1957

Bryn, F.

1 "Optimal Signal Processing of Three-Dimensional Arrays Operating on Gaussian Signals and Noise", JASA 34 (1962) 289-297

Courant, R. and Hilbert, D.

1 Methods of Mathematical Physics, Volume II: Partial Differential Equations, Interscience, New York, 1962

Cramer, H.

1 Mathematical Methods of Statistics, Princeton University Press, Princeton, 1963

Cron, B.F. and Sherman, C.H.

1 "Spatial-Correlation Functions for Various Noise Models", JASA 34 (1962) 1732-1736

2 "Addendum: Spatial-Correlation Functions for Various Noise Models (JASA 34, 1732-1736)", JASA 38 (1965) 885

Feldman, J.

1 "Equivalence and Perpendicularity of Gaussian Processes", Pacific J. Math. 8 (1958) 699-708

2 "Correction to 'Equivalence and Perpendicularity of Gaussian Processes' ", Pacific J. Math. 8 (1958) 1295-1296

Feldman, J.

3 "Some Classes of Equivalent Gaussian Processes on an Interval", Pacific J. Math. 10 (1960) 1211-1220

Gaarder, N.T.

1 The Design of Point Detector Arrays, Ph.D. Dissertation, Stanford University, 1965

2 "The Design of Point Detector Arrays, I", IEEE Trans. Info. Th. IT-13 (1967) 42-50

3 "The Design of Point Detector Arrays, II", IEEE Trans. Info. Th. IT-12 (1966) 112-120

Goode, B.

1 "Comments on 'Detection of Random Acoustic Signals by Receivers with Distributed Elements: Optimum Receiver Structures for Normal Signal and Noise Fields' ", JASA 39 (1966) 1193-1194

Grenander, U.

1 "Stochastic Processes and Statistical Inference", Ark. Mat. 1 (1950) 195-277

Hajek, J.

1 "On a Property of Normal Distributions of any Stochastic Process", Czech. Math. J. 8 (1958) 610-618 Translated in Amer. Math. Soc. Translations in Prob. and Stat. (1961)

Halmos, P.R.

1 Measure Theory, Van Nostrand, Princeton, 1950

Hamming, R.W.

1 Numerical Methods for Scientists and Engineers, McGraw-Hill, 1962

Hörmander, L.

1 Linear Partial Differential Operators, Academic Press, New York, 1964

Kailath, T.

1 "Correlation Detection of Signals Perturbed by a Random Channel", IRE Trans. Info. Th. 6 (1960) 361-366

Kakutani, S.

1 "On Equivalence of Infinite Product Measures", Ann. Math. 49 (1948) 214-224

Kallianpur, G. and Oodaira, H.

1 "The Equivalence and Singularity of Gaussian Measures", Chapter 19 in Time Series Analysis, Murray Rosenblatt, ed., Wiley, 1963

Kingman, J.F.C. and Taylor, S.J.

1 Introduction to Measure and Probability, University Press, Cambridge, 1966

Lavrentiev, M.M.

1 Some Improperly Posed Problems of Mathematical Physics, Springer-Verlag, New York, 1967

Liusternik, L.A. and Sobolev, V.J.

1 Elements of Functional Analysis, Frederick Ungar, New York, 1961

Lowenstein, C.D. and Anderson, V.C.

1 "Quick Characterization of the Directional Response of Point Array", JASA 43 (1968) 32-36

Martel, H.C. and Mathews, M.V.

1 "Further Results on the Detectability of Known Signals in Gaussian Noise", Bell Sys. Tech. J. 40 (1961) 423-451

Middleton, D. and Groginsky, H.L.

1 "Detection of Random Acoustic Signals by Receivers with Distributed Elements: Optimum Receiver Structures for Normal Signal and Noise Fields", JASA 38 (1965) 727-737

Morse, P.M. and Ingard, K.U.

1 Theoretical Acoustics, McGraw-Hill, New York, 1968

Officer, C.B.

1 Introduction to the Theory of Sound Transmission, with application to the ocean, McGraw-Hill, New York, 1958

Parthasarathy, K.R.

1 Probability Measures on Metric Spaces, Academic Press, New York, 1967

Rao, C.R. and Varadarajan, V.S.

1 "Discrimination of Gaussian Processes", Sankhya, Series A, 25 (1963) 303-330

Root, W.L.

1 "Singular Gaussian Measures in Detection Theory", Chapter 20 in Time Series Analysis, Murray Rosenblatt, ed., Wiley, 1963

Shearman, E.D.R.

1 "Non-collinear and Cylindrical Multiplicative Arrays", J. Brit. IRE, December 1963, 481-484

Shepp, L.A.

1 private communication

2 "The Singularity of Gaussian Measures in Function Space", Proc. Nat. Acad. of Sciences 52 (1964) 430-433

3 "Radon-Nikodym Derivatives of Gaussian Measures", Ann. Math. Stat. 37 (1966) 321-354

Schultheiss, P.M.

1 Degradation of Target Detectability Due to Clipping, Progress Report #6, General Dynamics/Electric Boat Research, Department of Engineering and Applied Science, Yale University, June 1963

2 Passive Detection of a Sonar Target in a Background of Ambient Noise and Interference from a Second Target, Progress Report #17, General Dynamics/Electric Boat Research, Department of Engineering and Applied Science, Yale University, September 1964

Skolnik, M.I., Nemhauser, G., and Sherman III, J.W.

1 "Dynamic Programming Applied to Unequally Spaced Arrays", IEEE Trans. Antennas and Propagation, January 1964, 35-43

Slepian, D.

1 "Some Comments on the Detection of Gaussian Signals in Gaussian Noise", IRE Trans. Info. Th. 4 (1958) 65-68

Sokolnikoff, I.S.

1 *Tensor Analysis*, second edition, Wiley, New York, 1964

Thomas, T.Y.

1 *Concepts from Tensor Analysis and Differential Geometry*, Academic Press, New York, 1965

Usher, T.

1 *Signal Detection by Arrays in Noise Fields with Local Variations*, Progress Report #2, General Dynamic/Electric Boat Research, Department of Engineering and Applied Science, Yale University, March 1963

Vanderkulk, W.

1 "Optimum Processing for Acoustic Arrays", J. Brit. IRE, October 1963, 285-292

Varberg, D.E.

1 "On Equivalence of Gaussian Measures", Pacific J. Math. 11 (1961) 751-762

Wilkinson, J.H.

1 *The Algebraic Eigenvalue Problem*, The Clarendon Press, Oxford, 1965

Yaglom, A.M.

1 "On the Equivalence and Perpendicularity of Two Gaussian Probability Measures in Function Space", Chapter 22 in *Time Series Analysis*, Murray Rosenblatt, ed., Wiley, New York, 1963

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) General Dynamics Corporation Electric Boat division Groton, Connecticut	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

PROCESSING OF DATA FROM SONAR SYSTEMS, VOLUME VII

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
July 1, 1968 to April 30, 1970

5. AUTHOR(S) (First name, middle initial, last name)

Franz B. Tuteur, John H. Chang, Verne H. MacDonald, and James P. Gray

6. REPORT DATE August 11, 1970	7a. TOTAL NO. OF PAGES 361	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. N00014-68-C-0392 b. PROJECT NO c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) U417-70-051 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10. DISTRIBUTION STATEMENT

Each transmittal of this document outside the agencies of the U.S. Government must have prior approval of the Office of Naval Research (Code 466)

11. SUPPLEMENTARY NOTES Reproduction of this publication in whole or in part is permitted for any purpose of the United States Government.	12. SPONSORING MILITARY ACTIVITY Office of Naval Research Washington, D.C.

13. ABSTRACT

Volume VII contains four progress reports (38, 39, 40, and 41), three of which are continuations of work covered in earlier volumes. These deal with the effects of anistropy of the background noise field, also referred to as interference noise. The three reports cover the form of the optimum detector (No. 38), the behavior of adaptive detectors (No. 39), and bearing estimation under these noise conditions (No. 40). Report No. 41, which deals with the effect of signal models and the various possibilities for singular detection, is a substantial departure from the work described in previous progress reports.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Antisubmarine Warfare						
Sonar						
Optimum Detector						
Adaptive Detector						
Signal Models						